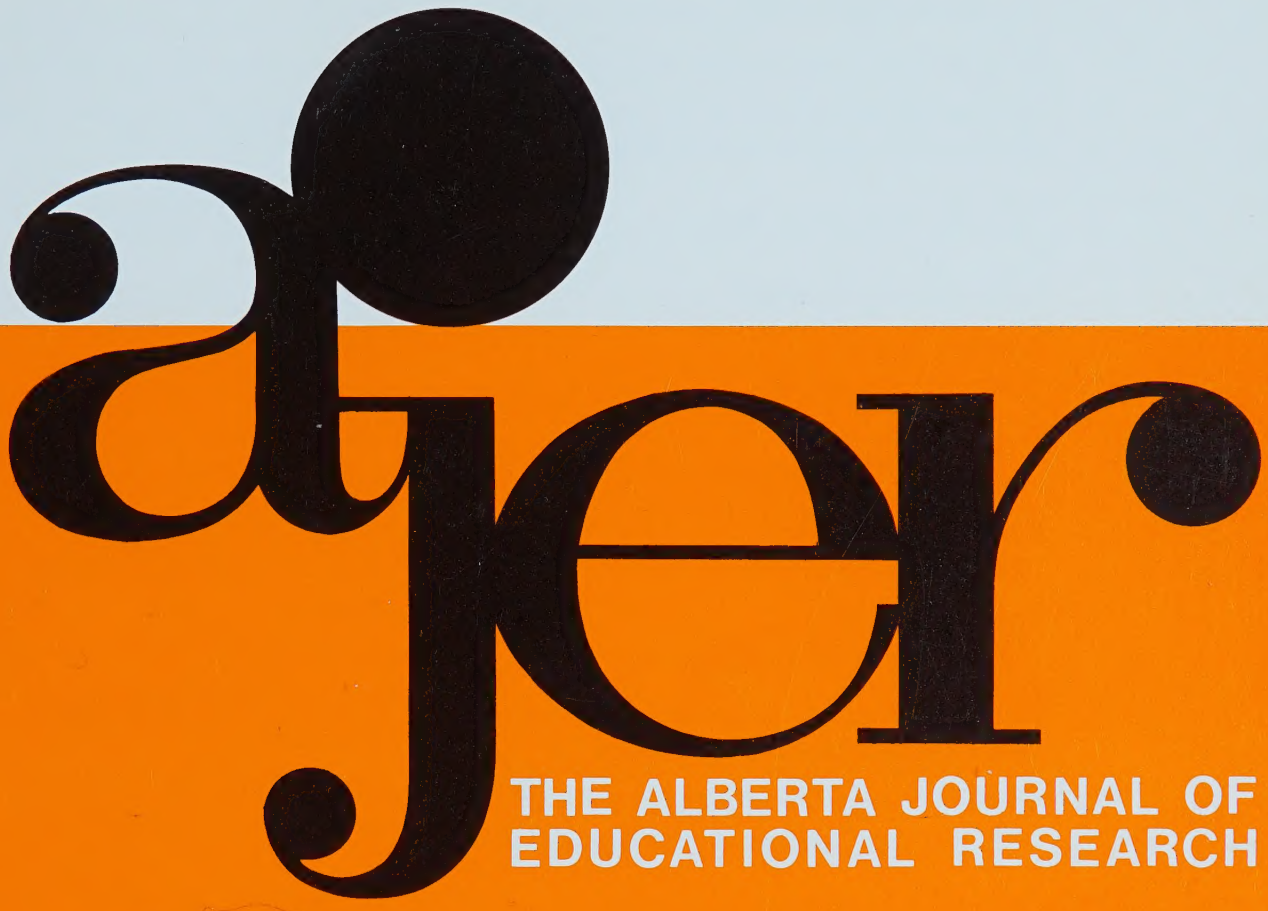


THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XXXVIII NUMBER 1 MARCH 1992

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

The Alberta Journal of Educational Research

Published in March, June, September, and December
by the Faculty of Education, University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation, and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Editor: Robert H. Short

Consulting Editors

James Chapman
Massey University

Rod Clifton
University of Manitoba

Dianne Common
Lakehead University

Carol Crealock
University of Western Ontario

John Hattie
University of Western Australia

Billie Housego
University of British Columbia

Linda Larocque
University of Alberta

Stewart J.H. McCann
University College of Cape Breton

Les McLean
Ontario Institute for Studies in Education

Linda Phillips
Memorial University of Newfoundland

Sharon Robertson
University of Calgary

Faculty Advisory Committee

John G. Paterson (Chair)
Fern Snart
Joyce Edwards
Graham Fishburne

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada, from the Alberta Advisory Committee for Educational Studies, and from the Office of Research Services, University of Alberta.

The subscription rate is $32.00 per year for individuals, $40.00 per year for institutions. Add $8.00 for delivery outside Canada. Single copies are $8.00 each. Subscriptions and sales in Canada will be charged 7% GST. Please make cheques payable to *The Alberta Journal of Educational Research*. Back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.

Publications Mail Registration Number 1436

Faculty of Education
University of Alberta

The Alberta Journal of Educational Research

Volume XXXVIII, Number 1, March 1992

Editorial	**1**	A Note of Thanks
Peter Evans, Lloy Falconer, Jocelyn Groves, Daniel Rubin, David Mather	**3**	An Evaluation of the IBM Principle of Alphabet Literacy System (PALS)
Hope-Arlene Fennell	**9**	An Investigation of the Relationships Between Organizational-Cultural Linkages and Teachers' Stages of Concern Toward a Policy Implementation
Bernard Davis, A.R. Bruce Clarke, James Francis, George Hughes, John MacMillan, John McNeil, Peter Westhaver	**27**	Dress for Respect: The Effect of Teacher Dress on Student Expectations of Deference Behavior
I.I. Dow, W.F. Oakley	**33**	School Effectiveness and Leadership
B.E.J. Housego	**49**	Monitoring Student Teachers' Feelings of Preparedness to Teach, Personal Teaching Efficacy, and Teaching Efficacy in a New Secondary Teacher Education Program
Joe Fris	**65**	Principals' Encounters with Conflict: Tactics They and Others Use

ISSN 0002-4805

Acknowledgment

The quality of *AJER* depends heavily on its manuscript review process. The critical and constructive reviews of the following people during 1991 have contributed enormously to the journal's success.

Andrews, J.
Bachor, D.G.
Bain, B.C.
Blakey, J.M.
Bowd, A.D.
Browne, M.P.
Butt, R.L.
Calder, P.
Campbell, D.S.
Carney, R.J.
Cathcart, W.G.
Chapman, J.W.
Clandinin, D.J.
Clifton, R.A.
Common, D.
Cousins, B.
Crowhurst, M.
Cummins, J.
Dawson, A.J.
Dodd, J.
Donald, J.G.
Fagan, W.T.
French, C.
Frender, R.H.
Gambell, T.J.
Goldberg, J.S.
Greene, M.L.
Hannay, L.
Hattie, J.
Hersom, N.L.
Hoge, R.G.
Holdaway, E.A.
Horowitz, M.
Housego, B.E.J.
Howe, N.
Hunka, S.M.
Ilott, H.G.
Jacknicke, K.G.
Kass, H.J.
Kelsey, J.G.T.
Kieren, T.E.
Kirby, J.R.
Larocque, L.
Leithwood, K.A.
Levin, M.A.
Lucas, B.
Magnusson, K.
Malicky, G.V.
Martin, R.A.
Maynes, W.
Mazurek, K.
McIntosh, R.G.
McKay, D.A.
McKinnon, D.
McLean, L.D.
McNay, M.L.
Meyer, J.R.
Millar, M.A.
Muir, J.A.
Muir, W.
Nagy, G.P.
O'Neill, G.P.
Oberg, A.A.
Pannu, R.
Patsula, P.
Patterson, R.S.
Phillips, L.
Ragsdale, R.G.
Rejskind, G.
Ricker, E.W.
Rogers, W.T.
Ryan, D.W.
Sackney, L.
Sawada, D.
Sheehan, N.M.
Sherrill, J.
Short, R.H.
Stewin, L.L.
Taafe, R.V.
Urion, C.A.
Violato, C.
Wilgosh, L.R.
Wood, B.A.
Wright, R.L.
Young, B.

Editorial

A Note of Thanks

At the beginning of each March issue of *AJER* the contribution of the past year's reviewers is acknowledged. One of the highlights that I have experienced since taking over as the editor of this journal is the positive and professional attitude of the individuals whom I have asked to review manuscripts. With very few exceptions they appear honored to be asked. As we all know there is very little tangible reward for reviewing what are sometimes very difficult manuscripts. It is just one more of the many things that we are asked to do each day in addition to our normal work load. However, it is my experience that most reviewers approach the task in a very constructive manner. They appear to see the process as a distinct way that they can help towards nurturing and directing research in education. Criticism is rarely anything but constructive and in most cases reviewers will go to considerable length to offer specific advice for improvement. It is on the basis of this very positive experience that I would like thank all those who have reviewed for the journal in 1991.

I would also like to take this opportunity to acknowledge the contribution of several retiring members of *AJER*'s Consulting Editorial Board. Janet Donald, Lilian Katz, John McLeod, Lauran Sandals, and Keith Sullivan have served on the Board for three and, in some cases, four years. Their contribution in offering guidance and support to the editor, in finding appropriate and reliable reviewers, and in reviewing articles themselves has been greatly appreciated by the journal. In addition I would like to welcome four new Board members: James Chapman of Massey University, New Zealand, Linda Larocque of the University of Alberta, Stewart McCann of University College of Cape Breton, and Sharon Robertson of the University of Calgary.

Beginning with this issue *AJER* will provide both an English and a French version of the abstract of each article that it publishes. This in my mind is an addition that is long overdue. It is perhaps fitting that the journal is adopting this policy as the December 1992 issue will be devoted to educational research in Quebec. Colette Dufresne-Tassé and Jacques Desautels of the Université de Montréal are serving as guest editors for that issue.

RHS

Digitized by the Internet Archive
in 2024 with funding from
University of Alberta Library

https://archive.org/details/ajer_1992

Alberta Journal of Educational Research *Vol. XXXVIII, No. 1, 1992, 3-8*

Peter Evans
Lloy Falconer
Jocelyn Groves
Daniel Rubin

and

David Mather
University of Victoria

An Evaluation of the IBM Principle of Alphabet Literacy System (PALS)

PALS (principle of Alphabet Literacy System) is an IBM computer-based reading instruction program designed to develop reading and writing skills in the below-grade-6 level. It became available from IBM in 1987 and in the same year was evaluated in the Saanich School District of British Columbia. Adults with a wide diversity of educational and cultural backgrounds attended the program for a period of 100 hours spread over a 20-week period. Results showed significant reading gains for adults with beginning reading levels ranging from grade 2 to 7. Participant feedback emphasized the importance of skilled and empathic support from instructional staff.

Le PALS (Principle of Alphabet Literacy System) *est un didacticiel IBM de lecture conçu pour développer les habiletés de lecture et d'écriture des personnes de niveaux inférieurs à la sixième année. IBM le mit à la disposition du public en 1987; la même année, il fut évalué dans le District Scolaire de Saanich en Colombie Britannique. Des adultes de niveaux d'éducation très divers et issus de milieux culturels disparates participèrent au programme pour une durée de 100 heures s'étalant sur une période de 20 semaines. Les résultats ont révélé des gains significatifs en lecture pour des adultes à des niveaux initiaux de lecture allant de la 2ème à la 7ème année. Le feedback donné par les participants a mis l'accent sur l'importance de l'assistance experte et soutenue du personnel enseignant.*

Introduction

Illiteracy is a pervasive problem in Canada. A recent survey by Southam News (Nesbitt, 1987) revealed that 24% of Canadian residents 18 years and older are illiterate. The study also found that approximately 30% of Canadian children

Peter Evans, a professor of language arts, has published numerous articles on aspects of language learning and collaborated in the writing of student texts for the junior high school reading program.
Lloy Falconer received her PhD and is currently working for School District No. 63 (Saanich) in British Columbia.
Jocelyn Garwood (formerly Groves) received her MA in 1990. She is currently working in the Educational Technology Centre.
Dan Rubin received his MA and is currently teaching in the school district.
David Mather is currently studying for his PhD.

do not complete high school, and even among those who do 7% lack sufficient reading and writing skills to function successfully in society.

Illiteracy limits an individual's job opportunities (Nickerson, 1985), reduces self-esteem (Bowren, 1987), and results in an overall reduction of productivity in society. The Southam Literacy survey (Nesbitt, 1987) estimated its economic cost to Canada at two billion dollars per year. For reasons such as these there has recently been a concentration of research on developing effective adult literacy programs (Balmuth, 1988; Bowren, 1987; Heathington, 1987; Nickerson, 1985; Padak & Padak, 1987; Roscow, 1988; Schlossberg, 1987; Schuman, 1989) including a focus on computer-assisted instruction (CAI) (Imel, 1988; Snyder & Herrlinger, 1986; Turner, 1988a, 1988b; Young & Irwin, 1988). This article reports on an evaluation of one such program implemented in 1988 under government funding in the Saanich School District near Victoria, British Columbia.

The program was designed by IBM and marketed in 1987 under the trade name of PALS (Principle of Alphabet Literacy System, IBM Canada Ltd., 1987). It is one of three major systems for purchase in adult literacy programs where the computer serves as the primary medium of instruction, with teachers acting as facilitators for the learning process (Turner, 1988a). The other two are Programmed Logic for Automatic Teaching Operations (PLATO) and Computer Curriculum Corporation (CCC). They differ in that PALS aims at grade levels 0 to 5, whereas both PLATO and CCC are designed for literacy levels of grade 4 or higher. At the time of this evaluation (1988), the PALS installation in Saanich was the first of its kind in Canada.

The researchers were commissioned by the funding agency to provide an evaluation of the effectiveness of the program on a group of adult volunteers characterized by wide variations in demographic variables. The diversity of needs in this population, as well as the absence of empirical evidence supporting the instructional efficacy of CAI with adult learners (Bean, Given, & Johnson, 1988) prompted the use of both qualitative and quantitative methods of assessment (Cronbach, 1975; Guba, 1987).

Method

Subjects

Because of the newness of the PALS program (IBM Ltd., 1987), no empirical data were available on its suitability for adult illiterates. The 1988 Saanich, BC installation was the first of its kind in Canada. Therefore, a wide range of adult subjects was sought from the general public and a variety of community agencies. Ages ranged from mid-teens to mid-60s and included a broad ethnic mix. Educational achievement ranged from completion of elementary school to, in the case of one immigrant, completion of university. The ratio of males to females was approximately 2:1. A total of 27 subjects completed the program.

Materials

The computer laboratory was furnished with four IBM InfoWindow touch screens and eight IBM personal computers. Hardware and software costs were approximately $120,000. Although the equipment could have been installed in a regular classroom, other considerations necessitated building a trailer-type laboratory at an additional cost of approximately $50,000. The new building

provided physical surroundings that were efficient and orderly, but also colorful and relaxed. The work stations were laid out in a manner that facilitated staff handling of individual problems without interfering with ongoing work at other stations. The facility was designed to handle 16 students at one time.

PALS (Principle of Alphabet Literacy System) is designed to develop reading and writing skills in a multimedia format. It attempts to connect speech with written communication using as a vehicle for learning a fable describing the invention of the alphabet. The fable is presented in comic-style format with audio accompaniment to the printed text. As an aid in translating speech sounds to print, visual demonstrations of the articulation of basic phonemes can be called up as required. The emphasis is on hearing the sounds in words, using alphabet symbols to represent the sounds, putting these symbols together to form words, and finally writing sentences. There is no insistence on correct spelling at the outset. Students are encouraged to spell words the way they hear them, as the computer is programmed to accept "sound-spelling." The focus is on word comprehension, word formation, and word usage in sentence building. The system is supplemented by computerized touch-typing and word processing programs.

Procedure and Design

Out of an original enrolment of 48, 27 subjects completed the course. Some students were unable to attend from the beginning while others dropped out as a result of conflicting work and family commitments. The remaining participants attended in groups of 7, 9, and 11 members during three separate one and a quarter-hour sessions. These took place in the afternoon and evenings on four days of each week over a period of 20 weeks.

The Woodcock Reading Mastery Test (Woodcock, 1973, Forms A and B) was administered at the beginning and end of the sessions to evaluate changes in reading skills. The Word Attack subtest of this battery served as a measure of sound-based word decoding skill. In addition, individual interviews were carried out to determine qualitative reactions to both content and administration of the PALS program.

The evaluation procedure emphasized both process and product evaluation, that is, the evaluators observed and interacted with the staff during the training period and also measured literacy outcomes on completion of the program. The procedure was comparable to the judgmental side of Stake's (1967) Countenance evaluative model. This model directs attention to antecedent conditions, transactions (e.g., training), and outcomes, breaking them down into planned and actual outcomes on which judgments can be made in relation to acceptable standards of performance.

The evaluators did not participate in the planning stage of the project as its one-of-a-kind nature, driven by the availability of new technology, led the agency to initiate evaluation after the installation was in place. Thus the emphasis was on determining whether the training setting (i.e., laboratory, equipment, and staff) was successful in stimulating effort in a diverse population. Without this knowledge, it would not have been possible to know whether the reading outcome measures were confounded by factors unrelated to the program design such as insufficient equipment or staff.

Two teachers certificated in traditional educational instruction served as facilitators. They had worked through the software themselves before using it with students. Both teachers were positive about the use of computer-assisted instruction as an educational medium. They were also aware of the need to provide a considerable amount of personal assistance during the initial stages of the program in order to develop confidence in students with little previous experience in using computers.

Results and Discussion

As the PALS program is aimed at a specific grade level range (0 to 5), assessment of treatment effects were compared against beginning Woodcock Mastery Reading Test (WMRT) levels. On this basis, of the 27 subjects who completed all sessions, six began at grades levels 1 to 2; 10 at grade levels 3 to 5 and 11 at grade levels 5 to 7. Table 1 indicates the grade improvement for each of the three groups. Paired t-test comparisons showed the progress for each group to be statistically significant.

Group I showed the smallest increase in grade level as well as the least variation in individual improvement. In contrast, both groups II and III showed considerably higher grade increases (three and four times the Group I increase respectively) and also much wider individual variations. Thus, at least on the basis of the WMRT results, the PALS program benefits are not restricted to beginning grade levels of 5 or below.

Qualitative and quantitative analysis of individual data revealed several reasons for variations in individual progress. These included absenteeism due to other commitments, differences in personal motivation, and differences in response to the phonological emphasis of the PALS program. The latter effect is indicated by Figure 1, which illustrates that in general the large reading gains were made by those subjects who were strong in sound-based word attack strategies. Three exceptions to this trend evidenced histories of specific reading difficulties. These subjects achieved ceiling levels in word-attack skills, but progressed less than three grades on the WMRT overall reading evaluation.

The participants of the program were generally unanimous in their liking for the nonstressful, self-paced learning environment provided by computer-assisted instruction, and equally impressed by the empathic one-on-one assistance provided by the instructors. The pupil-instructor ratio of 5:1 was influential in overcoming previous negative experiences, and was clearly a

Table 1
Woodcock Reading Mastery Test:
Mean Grade Level Improvements Versus Beginning Reading Level

Group	*Starting grade level*	*N*	*Mean grade level improvement*	*SD*	*t value*	*df*	*P(2-tail)*
I	1 - 2	6	.917	.412	5.46	5	.003
II	3 - 5	10	3.010	2.555	3.73	9	.005
III	5 - 7	11	4.236	1.803	7.79	10	.000

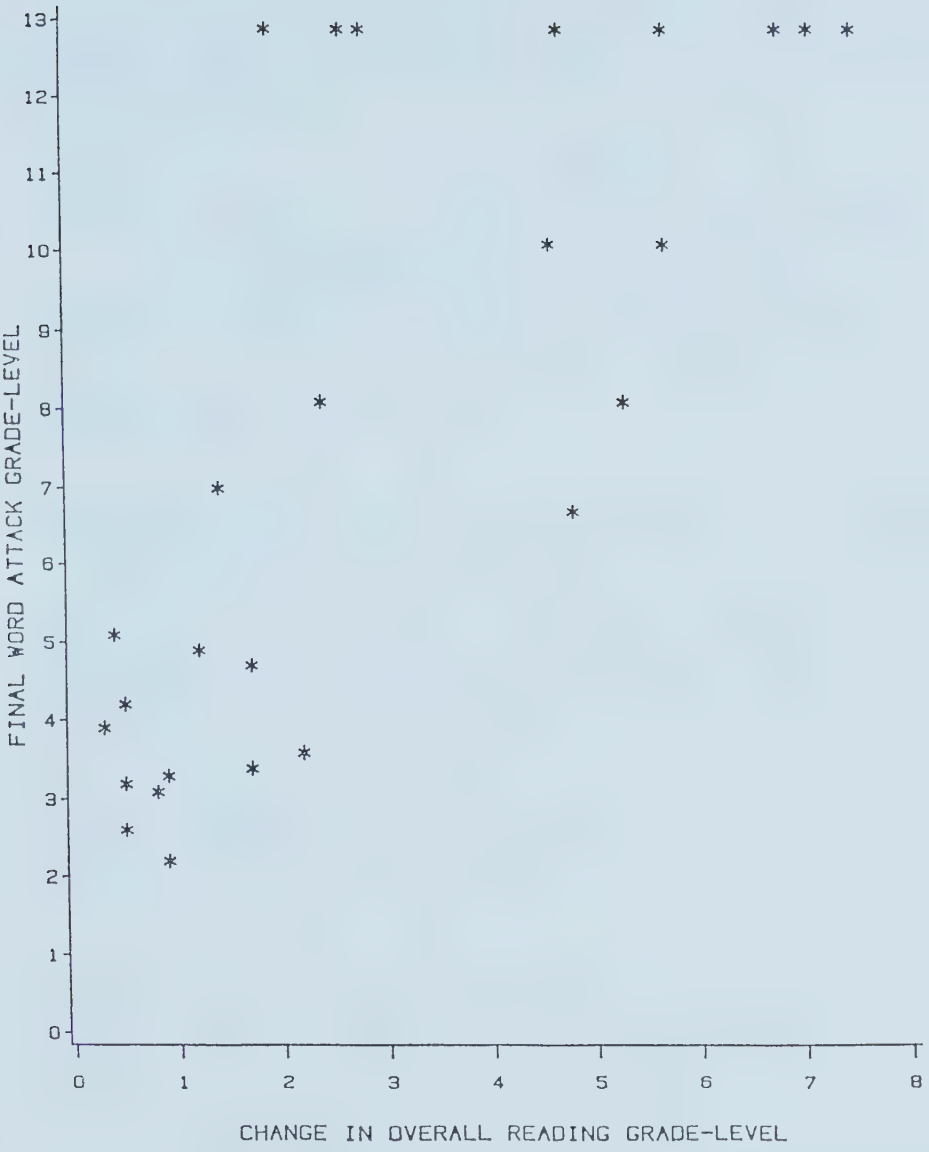


Figure 1. Woodcock Reading Master Test: Reading improvement versus final word-attack grade-level.

source of program success. Experience indicated that 12 rather than 16 students was an optimum class size for this facility.

Summary and Recommendations

The IBM Principle of Alphabetic Literacy System (PALS) produced significant reading gains in subjects varying widely in age, educational level, and cultural background. The Woodcock Reading Mastery Test (WRMT, Woodcock, 1973) scores calculated before and after an average of 100 hours instruction spread over 20 weeks found reading gains ranging from 0.4 to 7.5 grade levels. A comparison of individual progress with respect to beginning reading levels indicated that the most impressive gains were made by subjects possessing a minimum of a grade 2 reading ability as measured by the WRMT. It was generally indicated that subjects with the poorest sound-based word-analysis skills, as measured by the Word Attack subtest of the WRMT, made the least overall reading gains. Individual subject interviews suggested that the impor-

tance of readily available empathic assistance by well-trained instructors may be essential to a successful PALS implementation.

It is recommended that further evaluations of PALS should attempt to assess additional needs of participants who have deficient phonological skills.

References

Balmuth, M. (1988). Recruitment and retention in adult basic education: What does the research say? *Journal of Reading, 31*, 620-623.

Bean, R., Given, C., & Johnson, R. (1988). An evaluation of the implementation of microcomputers in an adult literacy program [Summary] *Adult Literacy and Technology National Conference* (Pittsburgh, PA, July 27-31, 1988). Rochester, NY: Gannett Foundations. (ERIC Document Reproduction Service No. ED 298 291)

Bowren, F.F. (1987). Adult reading needs adult research models. *Journal of Reading, 31*, 208-212.

Cronbach, L.J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist, 30*, 116-127.

Guba, E. (1987). Naturalistic evaluation. *New Directions for Program Evaluation, 34*, 23-43.

Heathington, B.S. (1987). Expanding the definition of literacy for adult remedial readers. *Journal of Reading, 31*, 213-217.

IBM Canada Ltd. (1987, March 6). IBM Infocourse: Principle of the Alphabet Literacy System (PALS). *Programming announcement*. Markham, ON.

Imel, S. (1988). *Computer-assisted instruction in adult literacy education.* Practice application brief. Washington, DC: Office of Educational Research and Improvement. (ERIC Document Reproduction Service No. ED 296 184)

Nesbitt, P. (1987). *Literacy in Canada—A research report* (Prepared for Southam News). Toronto, ON: Creative Research.

Nickerson, R.S. (1985). *Adult literacy and technology.* (Tech. Rep. No. 351). Champaign, IL: University of Illinois, Center for the Study of Reading. (ERIC Document Reproduction Service No. ED 266 420)

Padak, G.M., & Padak, N.D. (1987). Guidelines and a holistic method for adult basic reading programs. *Journal of Reading, 30*, 490-496.

Roscow, L.V. (1988). Adult illiterates offer unexpected cues into the reading process. *Journal of Reading, 32*, 120-124.

Schlossberg, N.K. (1987). Understanding and reading adult learners. *McGill Journal of Education, 22*, 9-18.

Schuman, R.B. (1989). Some assumptions about adult reading instruction. *Journal of Reading, 32*, 348-355.

Snyder, B., & Herrlinger, B. (1986). Computers are the answer to our adult-education needs. *American School Board Journal, 173*(9), 37.

Stake, R.E. (1967). The countenance of educational evaluation. *Teachers College Record, 68*(7), 523-540.

Turner, T. (1988a). An overview of computers in adult literacy programs. *Lifelong Learning, 11*(8), 9-12.

Turner, T. (1988b). Using the computer for adult literacy instruction. *Journal of Reading, 31*, 643-647.

Woodcock, R.W. (1973). *Woodcock Reading Mastery Tests.* Circle Pines, NM: American Guidance Service.

Young, D., & Irwin, M. (1988). Integrating computers into adult literacy programs. *Journal of Reading, 31*, 648-652.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 1, March 1992, 9-26*

Hope-Arlene Fennell
Lakehead University

An Investigation of the Relationships Between Organizational-Cultural Linkages and Teachers' Stages of Concern Toward a Policy Implementation

Educators and school organizations are faced with constant changes in educational policies related to curriculum and instruction. Implementing new policies presents many challenges to both educational administrators and teachers. Presented in this article are the findings from a recent study that investigated the existence of relationships between teachers' perceptions of the organizational-cultural linkages in the schools in which they taught and their stages of concern toward the implementation of a policy mandate. Other purposes included examining policy implementation strategies and investigating the patterns of organizational-cultural linkages that existed in the schools used as sites for the study. Key findings from this holistic study included the relationship of time factors to the concepts of top-down and bottom-up implementation strategies, differences in patterns of organizational-cultural linkages among the six schools, and the identification of relationships between categories of organizational-cultural linkages and teachers' responses and concerns to the implementation of a curriculum policy mandate, the Common Essential Learnings.

Les éducateurs et les organisations scolaires font face à des changements constants dans les politiques d'éducation concernant les programmes et l'instruction. La mise en pratique de nouvelles politiques présente plusieurs défis aussi bien aux administrateurs scolaires qu'aux enseignants. Dans cet article, on présente les résultats d'une étude récente dans laquelle on a examiné l'existence de relations entre les perceptions des enseignants concernant les liens organisationnels-culturels dans les écoles où ils enseignent et leurs niveaux d'intérêt pour la mise en pratique d'une politique mandatée. D'autres buts de l'étude étaient d'examiner les stratégies de mise en pratique des politiques et les patterns de liens organisationnels-culturels qui existaient dans les écoles qui ont servi à l'étude. Les résultats principaux de cette étude holistique comprenaient la relation entre les facteurs de temps et les concepts de stratégies de mise en pratique "de haut en bas" (top-down) et "de bas en haut" (bottom-up), les différences entre les patterns de liens organisationnels-culturels des six écoles, et l'identification de relations entre les catégories de liens organisationnels-culturels et les réactions et intérêts concernant la mise en pratique d'une politique mandatée de programme, les Apprentissages Communs Essentiels (Common Essential Learnings).

While the challenges facing educational organizations and the pressures weighing upon teachers increase in both number and complexity, results of research into change processes in education are fragmented and far from conclusive. Global issues such as our environment, economic reordering, and peace initiatives all have a sense of urgency about them. National leaders see

Hope-Arlene Fennell is an assistant professor in educational administration and curriculum and instruction in the School of Education. Her research interests include the development and implementation of curriculum and instructional policies, instructional leadership, and processes of teacher education.

schools as being crucial to the development of the human resources necessary to effectively contend with these and other issues. Most provinces are launching major initiatives in relation to schools despite their concerns with economic hardship and the increased public demands for accountability.

As leaders in education prepare to respond to environmental changes, they look to research for guidelines only to find that while much has been learned during the last decade the results appear to be fragmented. Some writers such as Corbett and Rossman (1989) highlight the importance of viewing changes as processes that are affected over time by the interaction of technical, political, and cultural factors. Others, such as Timar and Kirp (1987), have discussed dilemmas related to top-down and bottom-up implementation strategies. Top-down implementation strategies (Cuban, 1984; Seashore Louis & Dentler, 1988) refer to those strategies drafted at the top of school organizations to which organizational members must strictly adhere throughout implementation processes. Such strategies would include adherence to predetermined time lines and activities during implementation processes. Opposite top-down strategies are bottom-up strategies that are based solely on the discretion and actions of the individuals involved in implementation processes. Peters and Waterman (1982) contend that in order for change processes to be effective organizations have to be simultaneously loosely and tightly coupled. Simultaneous loose-tight coupling implies the coexistence of firm central direction with maximum individual autonomy for front line implementers of policies or innovations. For a number of years, Hall (1979) and his associates have advocated teachers' stages of concern as the basis for successful implementation. Research seems to suggest that if schools are to respond effectively to changes in their environment, implementation strategies must be based on societal and organizational factors as well as teachers' stages of concern.

The primary purpose of the study reported in this article was to investigate relationships between teachers' perceptions of the organizational-cultural linkages in the schools where they taught and their stages of concern about implementing provincially mandated instructional changes into their classroom teaching strategies. Other purposes were to describe the school division implementation strategy and to determine how it affected the implementation processes in each of the six schools examined in the research and to describe the patterns of organizational-cultural linkages noted in each of the six schools. The results of the study contribute to the literature by describing the variations in the patterns of linkages observed in six schools within one school system, by relating organizational linkages to teachers' concerns about implementation and by identifying timing as a key factor in understanding the relationships among the concepts of top-down, bottom-up, and simultaneous loose-tight coupling. In the next section, a particularly interesting innovation entitled the Common Essential Learnings is described. The Common Essential Learnings were the focus of the implementation processes for the research described in this article. Following the discussion of the innovation, the conceptual framework is presented and research methods are outlined. Findings are then summarized and discussed.

The Innovation: The Common Essential Learnings

Since the early 1980s, extensive planning for changes in curricula and instruction has been under way in Saskatchewan. The planning has resulted in the development of a policy statement centered around the Core Curriculum concept and has identified Required Areas of Study and a set of Common Essential Learnings (Saskatchewan Education, 1987; 1988a; 1988b). Consisting of six complex sets of skills, abilities, attitudes and knowledge, the Common Essential Learnings are to be developed throughout all areas of the curriculum. Communication, Numeracy, Creative and Critical Thinking, Technological Literacy, Personal and Social Values and Skills, and Independent Learning are the six areas that comprise the Common Essential Learnings.

The area of Communication focuses on the development and improvement of students' language skills throughout the curriculum and involves skills in the areas of reading, writing, listening, and speaking. Numeracy is concerned with the development of knowledge, skills, and appreciations of mathematical ideas, techniques, and application. The main rationale for numeracy is based on the increased need for individuals to interpret quantitative information as the use of technology increases in our society. Creative and critical thinking is intended to enhance the development of students' abilities to create and evaluate ideas, processes, experiences, and objects. Technological literacy involves developing an understanding of the ways in which society is shaped by technology, understanding the conflicting assumptions and values surrounding technology, and having the necessary skills to make responsible decisions about the social issues related to technology. Personal and social values and skills are thought to contribute to the development of responsible, compassionate individuals and are based on the moral value of respect for all persons. Independent learning focuses on the creation of opportunities and experiences that are necessary to assist students in becoming capable, self-reliant, self-motivated, and life-long learners.

The main features of the Common Essential Learnings are presented in four ideas. First, the Common Essential Learnings are not separate subjects and will not be taught as individual courses; instead they are incorporated in the curriculum. Second, students will not be evaluated on their acquisition of the Common Essential Learnings independently of the curriculum. Third, many of the same skills, processes, attitudes and values are the same for each of the Common Essential Learnings. Last, because of the Common Essential Learnings, teachers have been encouraged to develop a broader understanding of the teaching-learning process and to use a variety of instructional strategies.

In summary, the Common Essential Learnings are six sets of processes and skills designed to form the basic underpinnings that will shape curriculum and instruction in Saskatchewan as the transition is made to the Core Curriculum policy (Saskatchewan Education, 1987). Implementation of the Common Essential Learnings will require teachers to use a variety of instructional strategies and to make some modifications in their planning for instruction. In the next section, the discussion focuses on the questions and conceptual framework that guided this research on the implementation of the Common Essential Learnings.

The Conceptual Framework

The study of any educational change process involves three main sets of concepts: those describing to the implementation processes, those describing the organizational context of the implementation process, and those describing the actors involved in the implementation of the innovation or the policy (Fullan, 1982; Seashore Louis & Dentler, 1988). The concepts that were linked to form the framework (Figure 1) for this study included policy implementation strategies, organizational-cultural linkages, and teachers' stages of concern about an innovation. Each of the three concepts appears as a separate continuum on the framework. The underlying concept relating to each continuum involves thinking about the implementation of innovations from the technical, political, and cultural perspectives. Each of the concepts from the conceptual framework is briefly discussed below along with an overview of the interrelationships among them. Discussion begins with the three perspectives on innovation because this is the unifying concept underlying the conceptual framework.

Three Perspectives on the Implementation of an Innovation

House (1979; 1981), Tichy (1980) and Corbett and Rossman (1989) contend that the technical, political, and cultural perspectives of organizations need to be considered as part of the process of implementing an innovation. House (1981) indicates that while the technical perspective is still the most widely used approach to innovation, a shift in perspectives "from the innovation, to the innovation in context, to the context itself" (p. 28) has occurred. House also

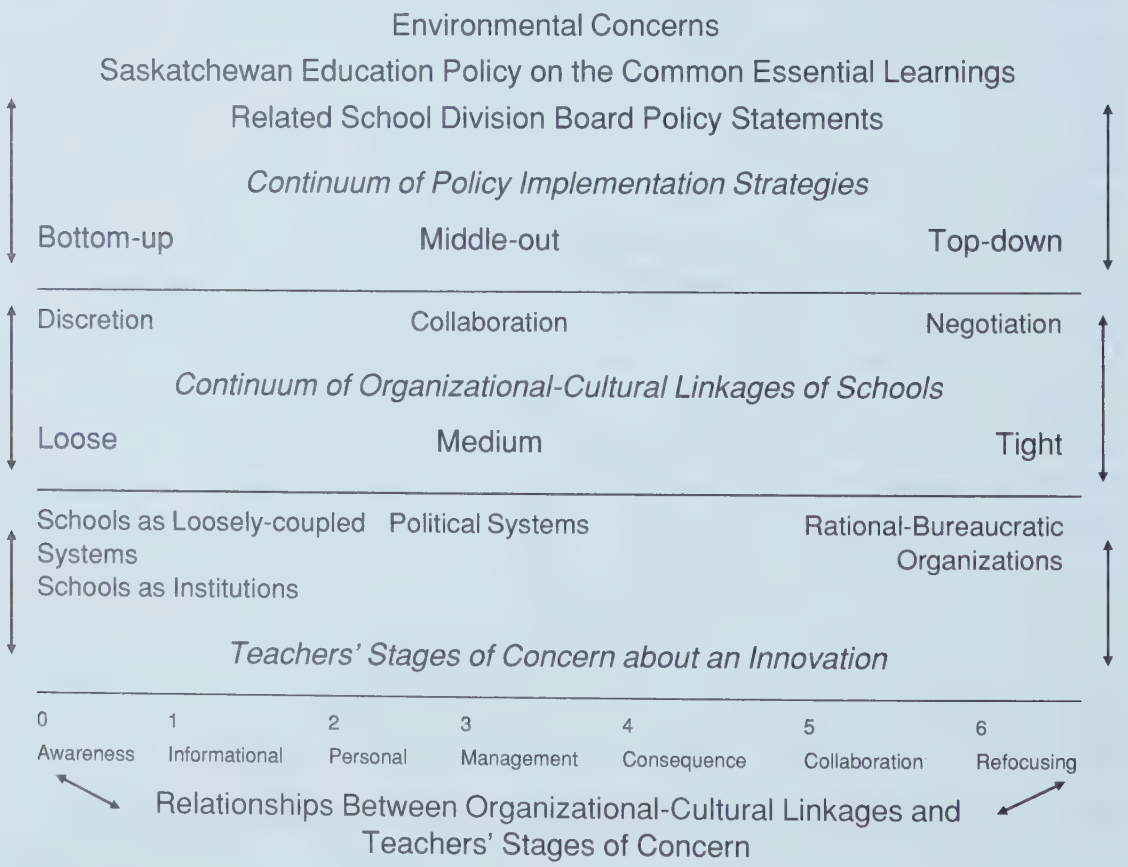


Figure 1. A framework relating policy implementation strategy, organizational-cultural linkages, and teachers' stages of concern about an innovation.

indicates that although each of the perspectives provides some important ideas for understanding the implementation process, the three perspectives in combination provide a more complete picture than any one alone. Tichy asserts that the best way to manage change within organizations is to assess the technological, political, and cultural cycles in organizations to determine how each will affect and be affected by innovations or change processes. Corbett and Rossman think that the three perspectives are present and active in all implementation processes. They note that problems with innovations that appear to be technical are also likely to have political and cultural ramifications. Corbett and Rossman (1989) conclude that, "At the school level, all three sets of processes can and did occur simultaneously" (p. 166). Each of the three perspectives is part of the policy implementation strategies, the organizational-cultural linkages, and the stages of concern expressed by teachers.

Policy Implementation Strategies

Strategies used in the implementation of educational policies have varied greatly. Wise (1983), Common (1983), Cuban (1984), Purkey and Smith (1985), and Timar and Kirp (1987) have discussed issues and dilemmas related to top-down and bottom-up implementation strategies. Barrett and Fudge (1981a) have discussed the importance of discretion and negotiation as part of the policy implementation strategy. Common (1983) and Seashore Louis and Dentler (1988) have discussed the importance of collaboration between policy implementers and the teacher-actors who will be expected to carry out the tasks introduced through the policy. Policy implementation strategies are likely to influence and to be influenced by the organizational context and the concerns felt by the actor-implementers.

Organizational-Cultural Linkages

The concept of the organizational contexts of schools and school organizations was defined and discussed using the term organizational-cultural linkages. Because of the varied definitions of the concept of coupling, Wilson and Corbett (1983) suggested the term *linkages* to describe the bonding components within organizations. Linkages were defined for this research as those mechanisms in schools that served to coordinate the activities of the people who worked there. Eight categories of linkages were identified (Wilson, Firestone, & Herriott, 1985), which included goal consensus among teachers, the facilitative leadership of the principal, communication among teachers and between teachers and the principal, the amount of conflict in the school, the level and order of student discipline in the school, and the degree to which the principal's influence on classroom instruction and on curriculum and resources was evident. Relating the eight categories of linkages to the concept of three perspectives on implementation, they were categorized as structural (technical), interpersonal (political), and cultural (cultural). Structural linkages included facilitative leadership and the principal's influence on classroom instruction and curriculum and resources. Interpersonal linkages included the amount of communication among teachers, between the principal and teachers, and the amount of conflict in the school. Goal consensus among teachers

and the level of student discipline in the school were categorized as cultural linkages.

The organizational-cultural linkages appear on the second continuum on the conceptual framework as part of three images of school organizations (Herriott & Firestone, 1984). Each image is described in terms of the looseness or tightness of the linkages that act as bonds to keep the organizational components together. The image of the rational-bureaucracy that clearly defined roles, rules, and delineated lines of authority was used to represent the most tightly linked organization. The images of loose coupling and institutions (Meyer & Rowan, 1983) were used to represent school organizations that were loosely linked, the opposite of rational bureaucracies. The image of school organizations as political systems was found between the loosely coupled and rational bureaucratic images. The political system image, which has characteristics of the other two images, focuses on the aggregation and dispersal of power throughout the school. Therefore, the political image represents school organizations that exist as negotiated orders that are constantly shifting to pressures from within and without the organization (Patterson, Purkey, & Parker, 1987). Organizational culture is also suggested as a linkage mechanism in organizations. Peters and Waterman (1982) strongly contend that cultural bonds can act as powerful linkage mechanisms coordinating the work and activities of individuals in the organization.

In summary, organizational-cultural linkages have been identified as mechanisms that serve to coordinate the activities of individuals in organizations and the bonds that link organizational components. Different patterns of linkages are displayed on a continuum as three images of school organizations. In the next section of the discussion, the concerns of teachers, who create a strong facet of the organizational context, are discussed.

Teachers' Stages of Concern about an Innovation

The third continuum on the conceptual framework represents teachers' stages of concern. Hall and Loucks (1978) contend that prior to and during any implementation process teachers go through a series of psychological stages that have been conceptualized as stages of concern toward the innovation. Hall (1979) indicates that the concerns that have been conceptualized in his work represent the motivations, perceptions, attitudes, and feelings that teachers experience in relation to implementing an innovation. The seven stages of concern have been categorized into three groups beginning with self-concerns, moving into concerns about managing the tasks of the innovation, and then becoming concerned with the impact of the innovation on the growth and development of their students. Hord, Rutherford, Huling-Austin, and Hall (1987) contend that teachers' concerns about innovations are developmental in nature and vary in intensity as the implementation process progresses.

Summary

The conceptual framework for the research was based on four concepts that appeared to be interrelated during the implementation process. The first was that, in implementation processes, the technical, political, and cultural aspects in all areas of the organization need to be addressed throughout the process. The second concept dealt with a continuum of implementation strategies,

ranging from top-down to bottom-up. This continuum was linked to the investigation of implementation strategies used for the implementation of the Common Essential Learnings in each of the schools. A third concept was related to the investigation of patterns of organizational-cultural linkages found in schools and was represented by a continuum of images of schools that ranged from the rational-bureaucratic, with tight linkages, to schools as loosely coupled systems or institutions, which are characterized by loose linkages. The last continuum on the framework represented teachers' stages of concern toward an innovation. The concerns are thought to be developmental in nature and to move from self concerns, through task concerns to concerns for the growth and development of students. The concerns continuum, in relation to the linkage continuum, addressed the main purpose of the research, which was to investigate the existence of relationships between teachers' stages of concern toward the implementation of the Common Essential Learnings and their perceptions of the organizational-cultural linkages in the schools where they taught. In the next section, selection of the research site, the methods and procedures used to collect and analyze the data attached to the framework are discussed.

Methods

In order to investigate implementation strategies, organizational linkages, and teachers' concerns, the study had to be designed holistically. To address the main research question regarding the existence of relationships between teachers' perceptions of the organizational-cultural linkages in their schools and their stages of concern toward the implementation of the Common Essential Learnings, the School Assessment Survey (Wilson et al., 1985) and the Stages of Concern Questionnaire (Hall & Loucks, 1978) were completed by each of the teachers. The contexts of the schools were studied to provide interpretative validity (Erickson, 1989) and to provide for the emergence of contextual factors that were affecting the implementation processes. Data for the study of the school contexts were collected through extensive observations, print materials, and interviews with principals, teachers, central office field staff, and central office administrators. The combination of both quantitative and qualitative data, as was used in this research, is supported by Wilson and Corbett (1983) and by Guba (1981). The emphasis on truth embedded in a social context is in keeping with critical research theory as described by Carr and Kemmis (1986). In the sections below, site selection, data collection, and data analysis procedures are described.

Selection of the Research Site

The site chosen for conducting the research was a large rural school division from which it was possible to choose six schools that had 10 or more teachers. Because the researcher was investigating the organizational-cultural linkages in schools, the schools from only one division were targeted to ensure that the goals of the central office administration were consistent. The researcher was also anxious that the six schools be selected from among elementary, kindergarten to grade 12, and secondary schools so that two of each were part of the sample. The sites were selected because it was believed that they were very

diverse in terms of the contexts of organization, principals' leadership styles, and the community contexts in which they were situated.

Data Collection

Four kinds of data were collected in answer to the questions of this study. Data were collected using two questionnaires, structured and unstructured interviews, field notes and observations, and printed materials related to the current and historical aspects of the implementation process. The data collection for the study took place in four phases over a four-month period. It is important to note, however, that the printed materials that were collected gave a historical view of the implementation that dated from two years prior to the beginning of this research process.

Phase one involved the selection of the research site and concluded with an initial visit to each of the selected schools, initial interviews with the Director of Education and the Associate Director, and the collection of historical documents related to the implementation. Phase two consisted of two visits, of at least one half day each, to each of the six schools. During phase two, data were collected through observations, field notes, questionnaires, interviews with principals and facilitating teachers, talking with parents, teachers, and students about the schools and communities, and through the collection of further printed materials about the implementation. Interviews with catalyst teachers were also conducted during phase two. Phase three involved the initial analysis and charting of this material. Further interviews with the Director and the Associate Director were also held. The SAS and the Stages of Concern Profiles for each school were also prepared, and data were linked to the cultural variables of the schools. Phase four involved the researcher in returning to each of the schools for further observation as a check on the original data and findings. Semistructured interviews were also held with each of the six school principals and with the Director and Associate Director. Following the last observations and interviews, the researcher made some final revisions and completed the analysis. Phase four took place almost two months after the completion of phase two, and immediately after completing phase three.

Pilot studies were conducted using the instruments prior to beginning the study in order to determine clarity of language, suitable administrative and scoring procedures, and which dimensions of the questionnaires were required to provide the necessary data for the research in a sample of schools that were similar in organization to those used in the actual study. Each is briefly described below.

The School Assessment Survey. The SAS was developed by Wilson et al. (1985) following intensive study of the concepts of coupling and linkages in school organizations. The SAS is a nine-dimensional questionnaire on which teachers' perceptions of the organizational characteristics of the schools where they teach are measured. When all teachers in a school are surveyed using the SAS, "a more accurate portrait can be produced, one that minimizes individual biases. Further, by systematically surveying all teachers in a given school, the full range of perceptions can be presented" (p. 2). For this study, eight of the nine dimensions were chosen: Goal consensus, facilitative leadership, vertical communication, horizontal communication, centralization of the

principal's influence on instruction, centralization of the principal's influence on curriculum and resources, staff conflict, and student discipline. These eight dimensions reflect the progression of relationships down the organizational hierarchy and represent the eight groups of organizational-cultural linkages related to the implementation of innovations (Wilson & Corbett, 1983). The SAS is composed of sixty-four items that comprise eight sets of linkages of eight items each.

The Stages of Concern Questionnaire. The SOCQ was developed from the Stages of Concern dimension of the Concerns-Based Adoption Model (Hall, 1979) and focuses on the concerns of individual teachers involved in implementation and change processes. Hall, George, and Rutherford (1979) indicate that the questionnaire was developed to provide an expedient measure of teachers' stages of concern regarding the implementation of innovations into their classroom teaching practices. Developed and validated over a three-year period, Hall et al. noted that after the being used in cross-sectional and longitudinal studies of 11 different implementation processes, the questionnaire accurately measures stages of concern about an innovation. The SOCQ is composed of 35 items that comprise seven scales of five items each.

Structured and semistructured interview schedules. In order to gather data about the specific policies of the school division with regard to the implementation of the Common Essential Learnings, separate structured interview schedules were prepared for the Director and Associate Director, the principals of the schools, teachers labeled as facilitating teachers, central office consultants, and catalyst teachers. During the last phase of the data collection, semistructured interviews were conducted with the Director and Associate Director and the school principals to gather their views and reactions to the preliminary findings of the research.

Field note observation data. Extensive field notes were made during the four observation visits to each of the six schools. The field notes were collected on the cultural and climate variables in each school.

Printed materials. Printed materials related to the implementation of the Common Essential Learnings were collected, beginning with the earliest planning stages and concluding with documents discussing current and future implementation plans. The materials included minutes of meetings, newsletters, and planning documents.

Data Analysis

The School Assessment Survey. Procedures for analysis were those given by the developers of the instrument. Scores for each school were determined by summarizing item scores for each of the eight dimensions for the teachers in each school. The school scores for each dimension were then placed on separate box and whisker charts where the scores for each dimension were compared with the scores normed by Wilson et al. (1985). On the charts, elementary school scores were compared with those normed for elementary schools, secondary schools were compared with scores normed on samples of secondary schools. Because of the span of the grade organization, scores for the two kindergarten to grade 12 schools were compared with the normed scores for both elementary and secondary schools.

The Stages of Concern Questionnaire. Individuals' responses on the SOCQ from each school were used to compute mean scores for the seven stages of concern for each of the six schools in accordance with instructions for the instrument. The mean scores were then translated to percentiles using the Stages of Concern Quick Scoring Sheet. The percentiles for each of the Stages of concern were then placed on a graph to create a profile of the teachers' Stages of Concern for each of the six schools. When the initial analysis was complete, the Stages of Concern profile and the profile on organizational-cultural linkages created from the SAS data were compared in a chart to determine the existence of relationships between the organizational contexts of each school and the teachers' stages of concern about implementing the Common Essential Learnings.

Structured and semistructured interview data. Data from both interview forms were transcribed separately. The data were coded initially based on whether the interviewees were principals, teachers, central office field staff, or central office administrators. The data were coded a second time to determine common themes related to the implementation processes. The data were coded a third time to determine the existence of data related to the organizational context of the schools, and a fourth time to determine the existence of data related more specifically to teachers' stages of concern toward the implementation process.

Field note data and printed materials. Field note data and printed materials were placed in categories based on the school from which each was collected and the chronological order of the collection. The field notes were then placed verbatim on the Cultural Variable Observation Form, a chart consisting of 14 cultural variables, in the category of the variable or variables they described or to which they were related. After the charts were completed for each of the visitations to the schools, the Cultural Variable Observation Forms were further analyzed to determine the examples that seemed to best represent or illustrate each of the variables on the charts. Some of the related printed materials were also included on the charts of cultural variables. Other printed materials were used as part of the description of the implementation strategies and the linkages between the implementation strategies and the organizational-cultural linkages of the schools.

Findings

School Division Implementation Strategy

The school division strategy for the implementation of the Common Essential Learnings was a mixture of top-down and bottom-up strategies, labeled as a *middle-out* strategy (Seashore Louis & Dentler, 1988). The top-down phase of the strategy included using the 1988-1989 school year as an awareness year, during which introductory inservice sessions were held for teams of principals and facilitating teachers from each of the schools, members of the central office staff, divisional and local boards of education, and, finally, all members of the teaching staff within the school division. Following the initial series of inservice sessions, macro-planning sessions were held to make provision for large-group inservice presentations and for funding to begin the actual implementation in the fall of 1989. The macro-planning sessions were led by the Director and Associate Director of Education, with input from principals

and teachers. The bottom-up phase of the strategy began in the fall of 1989. The in-school teams, comprised of the principal and facilitating teachers from each of the schools, were given the responsibility to work with the teachers in their schools to develop and carry out their own implementation plans. While the research was taking place, the implementation was taking place, to at least some degree, in each of the six schools.

Relationships Between School Division and School Implementation Strategies

In addition to strong leadership, extensive professional development, and clear expectations, the implementation strategy adopted at the school division level affected the subsequent strategies for implementing the Common Essential Learnings in each of the schools in two main ways. The first was that, although this implementation came about as a result of a provincial curriculum policy mandate, which was endorsed by the school division, the teachers and in-school administrators had opportunities to take part in how the implementation of the Common Essential Learnings would take place within their schools.

In addition to granting the teachers and principals a high degree of discretion for their in-school implementation strategies, the Director and Associate Director also provided strong support mechanisms to assist teachers with the implementation. The support mechanisms, which were categorized as cultural, political, and technical, ranged from such things as the Director of Education acting as a role model to positively reinforce the Common Essential Learnings, providing release time for teachers involved in planning for the implementation of the Common Essential Learnings, to providing the assistance of two catalyst teachers to go into classrooms and teach model lessons involving the various Common Essential Learnings.

Implementation Strategies for the Six Schools

The main aspects of the strategies for the implementation that were unique to each of the schools involved in the research were the degrees of discretion and collaboration that teachers had in implementing the Common Essential Learnings into their classroom teaching. In two of the schools, teachers indicated that they had had high degrees of discretion and that they had collaborated extensively among themselves and with the school administration in deciding on the timelines and strategies by which the implementation was to take place. Teachers in two other schools indicated lower degrees of discretion and less extensive collaboration as part of their implementation processes. In the remaining two schools, one group of teachers indicated that they had had moderate degrees of discretion and collaboration, whereas the teachers in the other school indicated that they had had low to moderate degrees of discretion and less extensive collaboration in deciding on the timing and methods for the implementation of the Common Essential Learnings. The variations in the degrees of discretion and collaboration noted were related to variations in the effects of the implementation processes on the reactions of the teachers and on the school organizations. The teachers who expressed the most positive responses and the least resistance to the implementation of the Common Essential Learnings were in schools where high degrees of discretion and collaboration were reportedly involved in the implementation processes.

The Patterns of Organizational-Cultural Linkages

The patterns of organizational-cultural linkages in the six schools were all different from each other and from the patterns observed by other researchers. While the two elementary schools appeared to have tighter linkages than the two secondary schools and one of the kindergarten to grade 12 schools, they had looser linkages than were previously associated with elementary schools (Herriott & Firestone, 1984; Wilson et al., 1985). The pattern for one of the elementary schools indicated tighter political and technical linkages and looser cultural linkages. Related to the continuum of organizational-cultural linkages on the conceptual framework, that school appeared to be close to the image of a negotiated political system. The image of the second elementary school appeared to be a synthesis of the rational-bureaucratic and the negotiated political system since the pattern of the second school appeared to have tighter technical linkages and looser political and cultural linkages.

The two secondary schools also represented two different organizational images. The pattern of linkages in the first school indicated tighter technical and political linkages and looser cultural linkages. Due to the degree of centralization of authority noted in the pattern, the image of the first school appeared on the continuum to be a combination of the institutional image and the image of a negotiated political system. The patterns for the other secondary school appeared more decentralized than that of the first. The cultural linkages appeared to be at mid-range, neither loose nor tight. The political linkages appeared to be looser than the norms of earlier researchers (Firestone & Herriott, 1981; Wilson et al., 1985) while the technical linkages appeared to be tighter than either the political or the cultural. The pattern for this school seemed to resemble that of a loosely coupled system.

One of the unique findings of this research was related to the patterns of organizational-cultural linkages of kindergarten to grade 12 (K-12) schools. Neither Firestone and Herriott (1981) or Wilson et al. (1985) indicate normed research for schools organized in this manner. Since there were no earlier norms on which to compare the patterns of K-12 schools, the patterns of the two schools were compared with each other. When compared, the patterns of the two K-12 schools appeared to be opposites. The pattern of one of the K-12 schools was very similar to the pattern that Wilson et al. had normed for secondary schools, with looser cultural, political, and technical linkages than the other. The image of this school appeared to be best represented as a combination of the loosely coupled system and the negotiated political system. The pattern for the other K-12 school indicated tighter linkages and appeared to be that of a rational bureaucracy with tighter cultural, political, and technical linkages than the patterns found within the norms for either elementary or secondary schools by Wilson et al. (1985).

Organizational-Cultural Linkages and Teachers' Stages of Concern

Despite the differences in patterns of both organizational-cultural linkages and teachers' stages of concern reported from each of the six schools, there did appear to be some interesting relationships between the two. In the two schools where there was a high degree of communication among the teachers, as indicated by the horizontal communication dimension on the SAS, and lower degrees of centralization of the principals' influence on instruction, less

teacher-principal communication about instruction, and less staff conflict, the teachers' refocusing concerns on the SOCQ were lower and they appeared to have more positive responses toward the implementation processes in their schools. Relating these findings to interview data on the implementation processes in the two schools, the principal of one of the schools indicated that, "We got input from teachers as to how fast they wanted to go in some areas." One of the teachers and a vice-principal also commented that, "We've been given a lot of leeway and so our staff has just gone with it. We could work at it at our own pace." Discussing the degree of communication among teachers with regard to the implementation, the vice-principal further commented that, "I see a lot of communication going on among teachers in various grade levels; there's been a lot of enthusiasm generated from that." The principal of the second school commented that, "Our strategy is being done cooperatively; we haven't been putting a lot of pressure on. I see teachers doing this on their own." He thought that by beginning the implementation more slowly and exerting less pressure on the teachers there was less resistance to the implementation. He commented that, "We have had next to no resistance to the implementation."

Discussion

Although a number of the findings from this research tend to support earlier findings of other researchers, others are less supportive or move beyond the findings of earlier researchers. The findings regarding the implementation strategy used by the school division the implementation of the Common Essential Learnings are similar to strategies described by Cuban (1984) and Seashore Louis and Dentler (1988). The strategy described in this research had both top-down and bottom-up components that created a middle-out strategy. The main strength of this strategy, to help create a vision and a feeling of ownership for the innovation among the innovators, appeared to be having the same type of positive results between the school division and the schools examined in this research process and those situations examined and discussed by earlier writers. However, these findings go beyond the work of earlier theorists by relating top-down and bottom-up aspects to time factors. The top-down strategies appeared to have been most useful in the initial phases of the implementation or wherever strong new spurts of initiative were or would be required. The bottom-up strategies appeared to be most effective in later stages of the process after the meaning of the innovation has become clearer to the teachers in the schools. Signals of the increased meaning and clarity of the innovations become evident when the goals for implementation become included in the instructional goals and objectives of classroom teachers.

Further support for the encouragement and development of bottom-up implementation strategies came from the findings of Barrett and Fudge (1981b), Common (1983), and Deasy (1984). They indicate that the implementation of policies and programs in schools is greatly assisted and enhanced by high degrees of collaboration among the teachers doing the implementing. High degrees of collaboration among teachers involved with implementation processes are likely to lead to greater understanding of the innovation and a

clearer meaning of what it will mean to their own work and that of their students.

Closely related to collaboration is the need for discretion among teacher-implementers. Collaboration and discretion are both positive and necessary factors to the success of implementation processes. While Common (1983) stresses the relationship between the strength of the political linkages controlled by teachers and the success of implementation processes, Deasy (1984) stresses the need for teacher discretion that encompasses the cultural and technical as well. The findings related to teacher discretion supported by cultural, political, and technical linkage mechanisms are also strongly supported by Corbett and Rossman (1989). They indicate that policy mandates such as the one for the Common Essential Learnings are most successfully implemented in situations where clear educational goals are enhanced by public and professional support mechanisms and where the mandate is feasible for the designated setting. Findings from this research relate to the earlier findings that the support provided through technical, political, and cultural linkages to each of the school implementation processes seemed to provide both the necessary vision and support to assist with the implementation.

The findings on the patterns of organizational-cultural linkages noted among the six schools corroborated some earlier findings, differed from others, and extended those of other researchers. Firestone and Herriott (1981) found that the varying degrees of linkages among the parts of organizations create different images of what schools are like and how they operate. Their research explored a continuum of organizational images on which schools have been examined and compared. Similar to their earlier findings, that individual schools could be seen as a combination of images, three of the schools in this research appeared on the continuum as combinations of images. Significant differences were noted between the findings of this research and the work of earlier researchers (Herriott & Firestone, 1984; Wilson et al., 1985) who attempted to relate patterns of linkages to the grade organizations of schools. The patterns of linkages reported among schools in this study appeared to be related more to the interactions and cultures of the schools than to the grade organizations. Although the two elementary schools appeared to have tighter linkages than the two secondary schools, they also appeared to have looser or fewer linkages than the elementary schools studied in earlier research by Herriott and Firestone and Wilson et al. Secondary schools also appeared to differ in patterns of linkages from the earlier findings of other researchers. Most significantly, the findings on patterns of linkages in K-12 schools make a unique contribution to the literature, since little or no research on schools of this nature appears to exist. The fact that the images of the two K-12 schools were opposites further substantiates that schools can be combinations of images and that each school is likely to have its own unique pattern of linkages that is likely to vary over time with changes in personnel, student enrollment, technical, or instructional policies.

The findings on relationships between dimensions of organizational-cultural linkages and teachers' stages of concern toward the implementation of the Common Essential Learnings add to the earlier theories and findings of

Barrett and Fudge (1981a), Common (1983), and Seashore Louis and Dentler (1988), who all stressed the need for teacher collaboration and discretion in planning for the implementation of curriculum policies and instructional changes into their daily work. The findings were also related to the concept of simultaneous loose-tight coupling noted by Peters and Waterman (1982), in that lower levels of direct supervision are replaced by higher levels of group collaboration and cultural influence. These findings go beyond the work of earlier theorists, however, in that they appear to identify specific categories and patterns of linkages that school organizations need to consider when attempting to develop a pattern of organizational-cultural linkages that will provide a nurturing environment for the implementation of innovations.

The findings noted in this research related to goal consensus and leadership in schools differed somewhat from the findings of earlier researchers. Both Ogawa and Bossert (1989) and Wilson and Corbett (1983) indicated that strong leadership and well-defined goals were important cultural linkages for assisting with implementation processes and assuring the spread of an innovation throughout a school organization. Despite these earlier claims, teachers in both of the schools where the implementation processes were proceeding most positively reported only moderate levels of goal consensus and a moderate degree of principal influence on their instructional work in the classroom. It is likely that higher degrees of collaboration assisted in promoting greater understanding among teachers so that they were able to work toward the common goals of implementation without high levels of consensus on sets of more general educational goals or high levels of facilitation and supervision by principals.

Summary and Implications

The purposes of this article were to describe the research design and findings from a recently completed study on the implementation of a curriculum policy mandate. The findings from this study regarding implementation strategies moved beyond those of earlier studies in that they attached a time element to the top-down and bottom-up sequence. The findings of this research also contributed with regard to patterns of organizational-cultural linkages in schools. Patterns of linkages in each of the six schools studied differed from each other and from those in the schools previously studied by other researchers. These unique patterns did not appear to be related to the grade organizations of the schools. There were no previous norms on which to compare the two kindergarten to grade 12 schools. In the relationships between the organizational-cultural linkages and their stages of concern about the implementation of the Common Essential Learnings, teachers indicated the fewest negative concerns and the least resistance to the innovation in the schools where they shared the greatest amount of communication and collaboration with their colleagues and where there was less direct centralization and control over their instructional work by principals. Only moderate degrees of consensus, toward a set of general educational goals, were noted in the two schools where teachers' responses to the implementation were most positive.

Implications for changes in the approaches to both research and practice need to be contemplated as a result of the findings of this study. Described in

this article are findings from six mini-implementation processes. Two of the six appeared to be moderately successful, two somewhat successful, and in the remaining two progress toward implementation was barely visible. In a school where little progress toward the implementation of the Common Essential Learnings was noted, one teacher indicated, "Many teachers feel top-downed. The idea that teachers have any kind of control isn't right." The teacher further indicated that many of the opportunities that teachers were given to have input into the implementation process was seen by at least some of the teachers as "pseudo control." Although more light has been shed on implementation processes and a certain degree of progress has been noted, one cannot quell the desire to search for more meaningful ways to develop and implement changes to curriculum and instruction. The implementation processes examined in this study appeared to be derived by a few individuals, and driven, front-end fashion, into the work lives of teachers as part of a larger provincial policy mandate.

As researchers and developers of curricular and instructional changes, perhaps we need to spend more time and resources on exploring questions about significant and desirable changes to curriculum and instruction with teachers, students, parents, and community members who are really the key players and beneficiaries of the outcomes of changes, both positive and negative. Perhaps instead of developing elaborate schemes for convincing individual teachers and participants to implement policies derived by relatively few individuals, all of the key players need to think deeply, together, about what changes are needed and how to bring them about. While thoughts of such dialogues may be fraught with some degree of messiness and discontinuity, the final results stand the chance of being no more expensive and at least as effective as innovations and schemes that teachers are presently being pressed to implement.

In 1982, Fullan noted:

> Where are we in the history of educational change? Surprisingly, given all the technological brouhaha of modernity, we are near the beginning, if we use serious attempts at *change in practice*—at implementation—as the criterion of progress. Being at the early stages of a complex undertaking, we have an obligation to be neither insultingly optimistic nor boringly pessimistic. We have seen many failures, but also some successes even under adverse circumstances. What is more, we understand better why things turn out as they do. Our knowledge is catching up to our information. Both are getting better. (p. 288)

Much of what is stated in Fullan's quotation could have been written today. We have seen failures in implementation processes, yet we have also seen successes. The findings related to top-down, middle-out, and bottom-up implementation strategies, the importance of teacher collaboration and discretion in implementation strategies, organizational-cultural linkages, and their relationships to teachers' responses and concerns about implementation processes will be useful in the development of future policies for changes to curriculum and instruction. The most important of these, collaboration, discussion, and discretion, are likely to be extremely useful for drawing all of the key participants into the development of future curricular and instructional policies that will truly bring about meaningful educational changes.

References

Barrett, S., & Fudge, C. (1981a). Examining the policy-action relationship. In S. Barrett & C. Fudge (Eds.), *Policy and action* (pp. 3-32). New York: Methuen.

Barrett, S., & Fudge, C. (1981b). Reconstructing the field of analysis. In S. Barrett & C. Fudge (Eds.), *Policy and action* (pp. 249-278). New York: Methuen.

Carr, W., & Kemmis, S. (1986). *Becoming critical: Education, knowledge and action research.* London: Falmer.

Common, D. (1983). Power: The missing concept in the dominant model of school change. *Theory Into Practice, 22*(3), 203-210.

Corbett, D., & Rossman, G. (1989). Three paths to implementing change: A research note. *Curriculum Inquiry, 19*(2), 163-190.

Cuban, L. (1984). Transforming the frog into a prince: Effective schools research, policy, and practice at the district level. *Harvard Educational Review, 54*(2), 129-151.

Deasy, R. (1984). *Guidelines for effective education policies. Educational Leadership, 42*(2), 5-10.

Erickson, F. (1989). *The meaning of validity in qualitative research.* Presentation to the American Educational Research Association Conference. Chicago: Teach 'em Inc.

Firestone, W., & Herriott, R. (1981). Images of organization and the promotion of educational change. In A. Kerchoff & R. Corwin (Eds.), *Research in sociology of education and socialization* (vol. 2, pp. 221-259). Greenwich, CT: JAI.

Fullan, M. (1982). *The meaning of educational change.* Toronto, ON: OISE Press.

Guba, E. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *ECTJ, 29*(2), 75-91.

Hall, G. (1979). The concern-based approach to facilitating change. *Educational Horizons, 57*(4), 202-208.

Hall, G., George, A., & Rutherford, W. (1979). *Measuring stages of concern about the innovation: A manual for the use of the SoC questionnaire.* Austin, TX: Research and Development Center for Teacher Education, University of Texas at Austin.

Hall, G., & Loucks, S. (1978). Teachers concerns as a basis for facilitating and personalizing staff development. *Teachers College Record, 80*(1), 36-53.

Herriott, R., & Firestone, W. (1984). Two images of schools as organizations: A refinement and elaboration. *Educational Administration Quarterly, 20*(4), 41-57.

Hord, S., Rutherford, W., Huling-Austin, L., & Hall, G. (1987). *Taking charge of change.* Alexandra, VA: Assocation for Supervision and Curriculum Development.

House, E. (1979). Technology versus craft: A ten year perspective on innovation. *Journal of Curriculum Studies, 11*(1), 1-15.

House, E. (1981). Three perspectives on innovation: Technological, political, and cultural. In R. Lehming & M. Kane (Eds.) *Improving schools—using what we know* (pp. 17-41). London: Sage.

Meyer, J., & Rowan, B. (1983). The structure of educational organizations. In J. Baldridge & T. Deal (Eds.), *The dynamics of organizational change in education* (pp. 60-91). Berkeley, CA: McCutchan.

Ogawa, R., & Bossert, S. (1989). *Leadership as an organizational quality.* Unpublished manuscript.

Patterson, J., Purkey, S., & Parker, J. (1987). *Productive school systems for a nonrational world.* Alexandria, VA: Association for Supervision and Curriculum Development.

Peters, T., & Waterman, R. (1982). *In search of excellence.* New York: Harper & Row.

Purkey, S. , & Smith, M. (1985). School reform: The district policy implications of the effective schools literature. *Elementary School Journal, 85*(3), 353-390.

Saskatchewan Education. (1987). *Policy direction for a core curriculum.*

Saskatchewan Education. (1988a). *Understanding the common essential learnings: A handbook for teachers.*

Saskatchewan Education. (1988b). *A resource kit for the implementation of the common essential learnings.*

Seashore Louis, K., & Dentler, R. (1988). Knowledge use and school improvement. *Curriculum Inquiry, 18*(1), 33-62.

Tichy, N. (1980). Problem cycles in organizations and the management of change. In J. Kimberly & R. Miles (Eds.), *The organizational life cycle* (pp. 164-183). San Francisco: Jossey-Bass.

Tichy, N. (1983). *Managing strategic change: Technical, political, and cultural dynamics.* Ann Arbor, MI: Wiley.

Timar, T., & Kirp, D. (1987). Educational reform and institutional competence. *Harvard Educational Review, 57*(3), 308-330.

Wilson, B. , & Corbett, H.D. (1983). Organization and change: The effects of school linkage on the quality of implementation. *Educational Administration Quarterly, 19*(4), 85-104.

Wilson, B., Firestone, W., & Herriott, R. (1985). *School assessment survey: A technical manual.* Philadelphia, PA: Research for Better Schools.

Wise, A. (1983). Why educational policies often fail: The hyperrationalization hypothesis. In J. Baldridge & T. Deal (Eds.), *The dynamics of organizational change in education* (pp. 93-113). Berkeley, CA: McCutchan.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 1, March 1992, 27-31*

Bernard Davis
A.R. Bruce Clarke
James Francis
George Hughes
John MacMillan
John McNeil

and

Peter Westhaver
Saint Mary's University

Dress for Respect: The Effect of Teacher Dress on Student Expectations of Deference Behavior

Although the concept of respect is commonly used by educators, it has not been defined for use in empirical studies. Although mode of dress has been shown to affect deference behavior, no general measure of deference has been defined. A 12-item questionnaire to measure respect for the teacher in a classroom setting was constructed using the method of summated ratings. The questionnaire was completed by 200 junior high school students who were shown either a photo of a teacher in formal dress or a photo of the same teacher in casual dress. The students expected more respect to be shown to the teacher in formal dress, t=2.982, p<.0032. *An item analysis showed differences in expectation to be significant only for those questions that identify deference behavior.*

Bien que le concept de respect soit d'usage courant parmi les éducateurs, son emploi dans les études empiriques n'a pas été défini. On a montré que la tenue a un effet sur le comportement de déférence; cependant aucune mesure générale de celui-ci n'a été définie. Un questionnaire à douze items pour mesurer le respect envers le maître en classe était construit par la méthode des indices additionnés (summated ratings). *Deux cents élèves du cycle secondaire moyen à qui on avait montré la photo d'un maître, soit en tenue formelle soit en tenue ordinaire, ont répondu au questionnaire. Les élèves ont prédit que le maître en tenue formelle sera plus respecté que le maître en tenue ordinaire* (t=2.982, p.<0032). *Une analyse d'items a révélé que les différences de prévisions étaient significatives seulement pour les items qui identifient le comportement de déférence.*

Teachers and principals often wonder whether their mode of dress affects the respect they get from their students. Published research, however, will not answer their question. Although educators use the concept of respect as a matter of course, this concept has not been defined for use in empirical studies.

Bernard Davis is an associate professor in the Faculty of Education.
A.R. Bruce Clarke, James Francis, George Hughes, John MacMillan, John McNeil, and Peter Westhaver are former students who now hold the Master of Arts in Education degree.

Although there is one rating scale for the communication of respect in interpersonal situations which requires an observer to identify what is said as communicating no, little, positive, deep, or the very deepest respect (Carkhuff, 1969), the most closely related research has measured deference behavior. In each study, the measure of deference used was part of the experimental design and followed experimental manipulation of subjects' behavior.

People more frequently violated pedestrian traffic signals if a person dressed as a high status individual first crossed against the light (Lefkowitz, Blake, & Mouton, 1955). Dimes left in phone booths were more likely to be returned if the person asking for the dime wore clothing indicating high social status (Bickman, 1971). Older subjects were more willing to answer a smartly dressed experimenter than an untidy experimenter (Lambert, 1972). Smartly dressed experimenters received more written feedback than casually dressed experimenters regardless of self-ascribed status (Giles & Chavasse, 1975).

Formal dress has been found to elicit more deferential behavior than casual dress in a university setting (Fortenberry, MacLean, Morris, & O'Connell, 1978). Subjects have responded more positively to conventionally dressed experimenters requesting signatures (Chaikin, Derlega, Yoder & Phillips, 1974; Samuel, 1972), cooperation (Raymond & Unger, 1972), or dimes (Klinke, 1977), handing out election information (Darley & Cooper, 1972), or requesting them to stop to respond to a survey (Walker, Harrison, & Costello, 1980), or to examine a petition (Bickman, 1974; Walsh, 1977).

Not all studies have shown positive response to high status dress. "Hippie" dress got more signatures than straight dress when a petition was presented to demonstrators at a peace march (Suedfeld, Bochner, & Matas, 1971). While the wearing of a tie increased positive response to an interviewer among middle class subjects, no difference was found for working class subjects (Green & Giles, 1973). Mothers were not more likely to accept a psychologist's recommendation if he wore a jacket and tie (Wasserman & Kassinove, 1976). A study done in an English suburb failed to find an effect for dress in a study in which different modes of dress were coupled with different accents (Giles & Farrar, 1979).

Some studies of the effect of dress have been done using photographs or drawings. Dress differences have a significant effect on the ratings given photos of adolescents by adolescent subjects (Hamid, 1969). Subjects rated drawings of a female job applicant as making a better impression when she was drawn as wearing a jacket or vest (Rucker, Taber, & Harrison, 1981). A woman in a jacket and skirt was seen as happier, more successful, more feminine, more attractive, more intelligent and more desirable as a friend than the same woman photographed in more casual dress (Harris et al., 1983).

No studies of the effect of dress on respect or deference behavior have been done in a classroom setting.

Instrument Development

As there is no measure of deference behavior suitable for use in a classroom setting, the experimenters designed an instrument using the method of summated ratings (Likert, 1932-33) that measured student expectation of behavior by fellow students that would indicate deference or respect. Relying on the observation that principal teachers generally get more respect from students

than substitute teachers, 22 questions concerning student behavior, which in the opinion of the experimenters indicated respect or lack of respect for the classroom teacher, were tested on a group of junior high school students, half of whom were asked "If the principal were teaching this class would your friends ...?" while the other half received the question "If a substitute teacher were teaching the class would your friends ...?" The 12 questions that survived this test, clearly differentiating between the principal and the substitute teacher, made up the final questionnaire. Of these questions, seven were phrased so that a positive answer indicated respect while five were phrased so that a positive answer indicated lack of respect.

Procedures

In the formal dress condition a picture of a male teacher of normal appearance (perhaps somewhat handsome) wearing a three-piece business suit was produced at the top of each questionnaire. In the casual dress condition, a picture of the same teacher wearing a T-shirt and jeans was at the top of the questionnaire. Pose, facial expression, and situation were typical for a classroom teacher and did not vary between the pictures.

The questionnaires were sorted into a random order and one was given to each of 200 grade 7 students in a junior high school in a predominantly middle-class suburb of Halifax by their home room teachers at the beginning of home room period. The students were instructed to answer the questions honestly and seriously and without looking at any other student's questionnaire.

Of the 200 responses 12 were discarded because of double checking of answers or other defects. The remaining 188 were scored and the mean scores of each group compared using Student's t-test for independent random samples (Student, 1908). The internal reliability of the instrument, Cronbach's alpha (Cronbach, 1951), was .79 for this sample.

Results

A significant difference was found, t=2.982, p<.0032, in favor of the formal dress condition. Means and standard deviations are shown in Table 1. The size of the effect found was measured using the standardized mean difference d_s (Cohen, 1977). The standardized mean difference was 0.4. Cohen categorizes a standardized mean difference of 0.5 as a medium size difference, "one large enough to be visible to the naked eye" (p. 26).

Table 1

	Casual dress	*Formal dress*
mean (over 12 questions)	33.6	36.6
standard deviation	7.32	6.55
N	96	92

t = 2.982
p<.0032
Means and standard deviations on respect questionnaire.

Table 2

If this teacher was teaching the class

Question	*Formal mean*	*Casual mean*	*Significance of Difference*
1. Would your friends do assignments given by him?	3.446	3.312	$p = .3643$, NSD
*2. Would your friends make smart remarks to him?	2.946	2.500	$p = .0056$
*3. Would your friends disrupt the class?	2.837	2.510	$p = .0458$
4. Would your friends pay attention in his class?	3.076	2.844	$p = .1311$, NSD
5. Would your friends think he is fair?	2.793	3.135	$p = .013$
6. Would your friends be quiet in his class?	2.543	2.167	$p = .0156$
*7. Would your friends be late for his classes?	3.457	3.000	$p = .0036$
8. Would your friends offer to help him if it seemed help was needed?	3.207	3.021	$p = .2747$, NSD
*9. Would your friends damage his classroom?	4.120	4.010	$p = .4632$, NSD
10. Would your friends talk to him about personal problems?	1.946	1.781	$p = .2497$, NSD
11. Would your friends work quietly in his class?	3.022	2.510	$p = .0015$
*12. Would your friends ignore detentions he gave?	3.207	2.781	$p = .0138$

Scale
1 definitely not
2 probably not
3 don't know
4 probably
5 definitely

* These questions were reflected before summing: for all questions a higher score implies a larger addition to the total score for perceived respect.

An item by item analysis of the results showed that all the questions except 5 discriminated in the direction anticipated. However, of these, only questions 2, 3, 6, 7, 11, and 12 discriminated strongly enough to reach significance as single items. These items measure features of respect that are features of deference behavior. Of those questions measuring what the experimenters considered other aspects of respect, 5, 8, and 10, two failed to reach significance, while question 5 was significant but in the wrong direction, students being more likely to expect their friends to think the casually dressed teacher fair.

A post hoc analysis showed the instrument consisting of the nine questions measuring expectation of deference behavior to differentiate between formal and casual dress at the p=.0006 level (t=3.50) with a standardized mean difference of 0.5 and an internal reliability of .80.

Conclusion

Formal dress appears to have an effect on junior high students comparable to the effect it has on adults. Students expected the formally dressed teacher to be shown more deference by their fellow students. This significant difference did not extend to expectations for behavior included in the authors' intuitive concept of respect which are not expectations for deference behavior. A possible practical implication of this research is that teachers for whom the first impressions of students are important (substitute teachers, student teachers, or any teacher on the first day of classes) might consider high status dress as

having the potential to increase the likelihood of student expectations of deference behavior.

References

Bickman, L. (1971). The effect of social status on the honesty of others. *Journal of Social Psychology, 85*, 87-98.

Bickman, L. (1974). Social roles and uniforms: Clothes make the person. *Psychology Today,* April 7 (11), 48-51.

Carkhuff, R. (1969). *Helping and human relations.* New York: Holt, Rinehart and Winston.

Chaikin, A.I., Derlega V.J., Yoder, J., & Phillips, D. (1974). The effects of appearance on compliance. *Journal of Social Psychology, 92*, 199-200.

Cohen, J. (1977). *Statistical power analysis for the behaviorial sciences* (rev. ed.). New York: Academic Press.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrica, 16,* 297-334.

Darley, J.M., & Cooper, J. (1972). The "clean for gene" phenomenon: The effects of student appearance on political campaigning. *Journal of Applied Social Psychology, 2*, 24-33.

Fortenberry, J.H., MacLean, J., Morris, P., & O'Connell, M. (1978). Mode of dress a perceptual cue to deference. *Journal of Social Psychology, 104*, 139-140.

Giles, H., & Chavasse, W. (1975). Communication length as a function of dress style and social status. *Perceptual and Motor Skills, 40*, 961-962.

Giles, H., & Farrar, K. (1979). The behavioral consequences of speech and dress styles. *British Journal of Social and Clinical Psychology, 18*, 109-210.

Green, W.P., & Giles, H. (1973). Reaction to a stranger as a function of dress style: The tie. *Perceptual and Motor Skills, 37*, 676.

Hamid, P.N. (1969). Changes in person perception as a function of dress. *Perceptual and Motor Skills, 29*, 191-194.

Harris, M., James, J., Chavez, J., Fuller, M.L., Kent, S., Massavari, C., Moore, C., & Walsh, F. (1983). Clothing, communication, compliance and choice. *Journal of Applied Social Psychology, 13*, 88-97.

Klinke, C.L. (1977). Effects of dress on compliance to requests in a field setting. *Journal of Social Psychology, 101*, 223-224.

Lambert, S. (1972). Reactions to a stranger as a function of style of dress. *Perceptual and Motor Skills, 35*, 711-712.

Lefkowitz, M., Blake, R.R., & Mouton, J.S. (1955). Status factors in pedestrian violations of traffic signals. *Journal of Abnormal and Social Psychology, 51*, 704-706.

Likert, R. (1932-33). A technique for the measurement of attitudes. *Archives of Psychology, 22*(140), 1-55.

Raymond, B.J., & Unger, R.K. (1972). The apparel oft proclaims the man: Cooperation with deviant and conventional youths. *Journal of Social Psychology, 87*, 75-82.

Rucker, M., Taber D., & Harrison, A. (1981). The effect of clothing variation on first impressions of female job applicants: What to wear when. *Social Behavior and Personality, 9*, 53-64.

Samuel, W. (1972). Response to bill of rights paraphrases as influenced by the hip or straight attire of the opinion solicitor. *Journal of Applied Social Psychology, 2*, 47-62.

Student (W.S. Gossett). (1908). The probable error of a mean. *Biometrica, 6*, 1-25.

Suedfeld, P., Bochner, S., & Matas, C. (1971). Petitioners' attire and petition signing by peace demonstrators: A field experiment. *Journal of Applied Social Psychology, 1*, 278-283.

Walker, M., Harrison, S., & Costello, S. (1980). The influence of appearance on compliance with a request. *Journal of Social Psychology, 112*, 159-160.

Walsh, E.J. (1977). Petition signing in town and on campus. *Journal of Social Psychology, 102,* 323-324.

Wasserman, T., & Kassinove, H. (1976). Effects of the type of recommendation, attire, and perceived expertise on parental compliance. *Journal of Social Psychology, 99*, 43-50.

Alberta Journal of Educational Research *Vol. XXXVIII, No. 1, March 1992, 33-47*

I.I. Dow
University of Ottawa

and

W.F. Oakley
Department of Education, Newfoundland and Labrador

School Effectiveness and Leadership

The literature on school effectiveness suggests that principal leadership is an essential ingredient in creating and maintaining an effective school. Research utilizing a contingency approach appears promising in studying the school-based phenomenon of school effectiveness. For the past two decades Fiedler's (1967) Contingency theory has been the most frequently used approach to study leadership from a situational perspective. A meta-analysis of studies employing the Fiedler model conducted by Crehan (1984) suggested that the theory may have limited applicability to schools. A study was designed to examine the effect of leadership style of the elementary principal on school effectiveness. Findings from the study would provide a test of Fiedler's theory in relation to a major concept in the literature, as well as test the concerns expressed by Crehan (1984) in her meta-analysis. The study measured school effectiveness as perceived by 900 teachers in 176 schools. Data were collected through a series of questionnaires. Leadership style of principals was measured by the Fiedler Contingency theory instruments; Least Preferred Co-worker (LPC), Group Atmosphere (GA), Task Structure rating scale (TS) and Principal Position power scale (PP). Two major hypotheses were tested. Results of the study challenge the suitability of Fiedler's Contingency theory in studying schools. Neither of the major hypotheses was supported. The major postulate of the theory that group effectiveness is determined by a combination of the leadership style of the principal and the situational favorability for the principal was not supported. The findings indicate that additional situational factors may be contributing significantly to group effectiveness. Major weaknesses in the measurement of leadership style in the Fiedler model are identified. Suggestions for modifications required to make the model more applicable to the school situation are presented.

La littérature ayant trait à l'efficacité de l'école suggère que le leadership du directeur est un ingrédient essentiel dans l'émergence et le maintien d'une école efficace. La recherche utilisant une approche de contingence paraît prometteuse pour ce qui est de l'étude du phénomène scolaire de l'efficacité de l'école. Durant ces deux dernières décennies, la théorie de contingence de Fiedler (1967) a été l'approche la plus utilisée dans l'étude du leadership du point de vue situationnel. Une méta-analyse des études utilisant le modèle de Fiedler conduite par Crehan (1984) a suggéré que la possibilité d'appliquer la théorie aux écoles serait peut-être limitée. Une étude était conçu pour examiner l'effet du style de leadership du directeur de l'école élémentaire sur l'efficacité de l'école. Les résultats de l'étude pourraient fournir une évaluation de la théorie de Fiedler concernant un concept important dans la littérature, et pourraient être un test des idées exprimées par Crehan (1984) dans sa méta-analyse. L'étude a mesuré l'efficacité de l'école telle que perçue par 900 enseignants dans 176 écoles. La collecte des données a été faite par une série de questionnaires. Le style de leadership des directeurs a été mesuré par les instruments de la théorie de contingence de

I.I. Dow, a former teacher and principal, is currently an associate professor of educational administration in the Faculty of Education.
Wayne F. Oakley has worked as a teacher, department head, curriculum consultant, central office administrator, and is presently Director of Curriculum and Instruction.

Fiedler: Least Preferred Co-worker (LPC), Group Atmosphere (GA), *l'échelle* Task Structure (TS), *et l'échelle* Principal Position power (PP). *Deux hypothèses majeures ont été testées. Les résultats de l'étude constituent un défi à la pertinence de la théorie de contingence de Fiedler dans l'étude des écoles. Aucune des hypothèses principales n'a été corroborée. Le postulat majeur de la théorie selon lequel l'efficacité du groupe est déterminée par un mélange du style de leadership du directeur et des conditions situationnelles favorables au directeur n'a pas été supporté. Les résultats indiquent que des facteurs situationnels additionnels contribueraient d'une manière significative à l'efficacité du groupe. Les principales lacunes dans la mesure du style de leadership dans le modèle de Fiedler sont identifiées. Des suggestions concernant les modifications requises pour rendre le modèle plus applicable à la situation scolaire sont présentées.*

Introduction

Researchers have been investigating the topic of organizational effectiveness for over 100 years (Lewin & Minton, 1986). In spite of the longevity of the construct and the popularity of the subject, a considerable amount remains to be done to arrive at a suitable model. The topic of school effectiveness has been receiving increased attention in the past few years. To date, the research on school effectiveness has identified a number of factors that appear important in identifying effective schools. One factor that appears consistently in all of the studies is principal leadership.

Early studies (Stogdill, 1948) in the area of leadership concentrated on the characteristics or traits that people brought to a particular position. Subsequent studies focused on the specific behaviors an incumbent displayed in a leadership position. In the last several years many researchers (Fleishman, 1973; House & Baetz, 1979) are suggesting that some form of contingency approach is necessary to determine leadership effectiveness. A number of contingency models are available in the literature; however, the theory proposed by Fiedler has received constant attention since it was presented in 1967.

Studies on school effectiveness have shown that a number of factors at the school level combine to produce an effective organization. Among these, the school principal has been suggested as playing a key role. The purpose of the present article is to examine the effect of leadership style of the principal on school effectiveness using Fiedler's Contingency theory.

School Effectiveness

Research studies on school effectiveness conducted during the 1950s and early 1960s, (Goodman, 1959; Mollenkopf & Melville, 1956) were largely optimistic in concluding that there was a direct relationship between the quality of school services provided to students and their subsequent achievement. However, Collard (1984) noted many methodological problems with the research at that time; for example, many of the studies did not examine the confounding influence of factors outside the school.

Consistent with this perspective, studies by Coleman et al. (1966) and Jencks et al. (1972) concluded that schools have little effect on a child's achievement that is independent of background and social context, and that variations in achievement scores are more likely to be a result of background factors rather than schooling. Such contradictory conclusions have contributed to the proliferation of research on school effectiveness during recent years.

The more recent literature on school effectiveness concludes that there are structural differences among schools that affect students' academic achievement. Studies on the determinants of achievement have been concerned with variables relating how schools and school districts are structured and make decisions, the process of change in schools and school districts, and the way in which classrooms and schools can increase the amount of time spent on productive instruction (Purkey & Smith, 1983, p. 428).

Weber (1971) was one of the first researchers to test the hypothesis that schools can make a difference. He studied four effective schools from a nomination list of 96 throughout the United States. He visited each of the schools, administered additional achievement tests, and interviewed and observed principals and staffs. From his case study reports on each school he found a number of factors common to each school, for example, high expectations, orderly climate, and stress on reading. Subsequent studies (Austin, 1978; Brookover & Lezotte, 1979; Shoemaker & Fraser, 1981) were designed to identify specific factors that differentiated between high achieving schools and low achieving schools. Factors associated with high achieving schools were as follows: positive pupil-teacher interaction, frequent informal classroom observations by the principal, a set of school-wide practices for reading instruction, attention to atmosphere conducive to learning, and open communication with parents and community.

A longitudinal study by Mortimore, Sammons, Stoll, Lewis, and Ecob (1986) not only concluded that schools do make a difference, but also identified a number of factors that promote effectiveness. These factors included purposeful leadership of the staff by the head teacher, involvement of the deputy head, involvement of teachers in the school's operation, consistency among teachers, structured sessions, intellectually challenging teaching, work centered environment, maximum communication between teachers and students, good record keeping, parental involvement, and positive climate.

A review of the literature conducted by Oakley (1989) delineated seven factors as having importance for school effectiveness, including: (a) strong leadership on the part of the principal, (b) high expectations by the principal and staff for student achievement, (c) a positive, orderly climate, (d) emphasis on basic skills development, (e) clearly stated, well-defined goals, (f) frequent and systematic evaluation of students, and (g) good home-school relations.

The Principal as Leader

A review of the literature on school effectiveness indicates that the principal plays an important leadership role. Wright and Renihan (1985) note that school leadership, that is, principal leadership, is critical in influencing how effective a school will be. There is a well-documented body of literature in educational administration (Fullan, 1982; Geering, 1980; Leithwood & Montgomery, 1981; Lipham, 1977) indicating the importance of the principal on many varied aspects of school operations. Rutter, Maughan, Mortimore, Quston, and Smith (1979) in an extensive longitudinal study of school effectiveness made the following observation:

> Having found evidence of considerable differences in both overall climates and the particular practices of schools, it is now important to go on to inquire how such climates become established and then maintained or changed. Obviously

> the influence of the head teacher (principal) is very considerable. We did not look in any detail at the particular styles of management and leadership which worked best; this is an issue which is now important to investigate. (pp. 202-203)

Further support for investigating the situational context of the principal comes from the work of Rowan, Bossert, and Dwyer (1983). They suggest the feasibility of using contingency theory to analyze the varying effects of organizational structures in school settings.

Contingency theories have been a dominant approach among leadership researchers for the last two decades (Schemerhorn, Hunt, & Osborn, 1982). The basic guiding principal of contingency theory is that the most appropriate leadership in terms of follower outcomes is a question of contingencies facing the leader (Hunt, 1984). As the study of leadership effectiveness evolved it has become increasingly apparent that leadership is a product of at least three components: the leader, the led, and the situation in which they all function (Crehan, 1984). Theory proposed by Fiedler (1964) has become one of the most widely used models in the study of the school principal (Crehan, 1984).

In summary, a number of researchers (Geering, 1980; Sackney, 1986) stress the pivotal role played by the principal in stimulating the performance of personnel in the school setting. Principal behavior appears to be influenced by a number of internal and external situational factors. These factors seem to have an impact on school effectiveness. Fiedler's contingency theory appears to be a sophisticated approach to examining the complex relationships among the leader, the led, and the situation. To date, Fiedler's contingency theory has not been utilized with regard to principal leadership and its effect on school effectiveness. The next section of the article describes the Fiedler model and its proposed strength in attempting to determine specific characteristics of school effectiveness.

Fiedler's Contingency Theory

Fiedler's theory maintains that different situations require different leadership styles. The leadership style of the leader and the degree of situational favorability of the group determines group effectiveness. Therefore, in schools, Fiedler (1967) suggests that group effectiveness is determined by a combination of the leadership style of the principal and the favorableness of the situation for the principal. This section of the article begins with a brief description of the situational favorableness dimension.

Fiedler (1967) developed a taxonomy of groups that included counteracting, coacting, and interacting. Definitions for each of the task groups identified were as follows: counteracting groups are typically engaged in negotiations and bargaining processes; coacting groups work relatively independently of other team members; and interacting groups need to work together to perform the task. Fiedler maintains that the leader's role must be different for each group type. He believed that there were three situational variables which were most important in the classification of interacting groups. Because the study reported in this article deals with interacting groups in a school situation, the three situational variables were measured. These included (a) the interpersonal relationship between leader and members; (b) the structure of the task; and (c) the leader's position power.

The group atmosphere variable is considered by Fiedler (1967) to be the most important of the three situational variables. It refers to the leader's effective relations with group members, to the acceptance the leader can obtain, and to the loyalty the leader can engender (Fiedler, 1967). Studies by Fishbein, Landy, and Hatch (1969) and Nebeker (1975) suggested that this dimension was the most important single determinant of leadership. Leader member relations is assessed by having either the leader or the group complete the Group Atmosphere Scale (GA). The instrument is a 10-item semantic differential scale. A further description of the instrument is provided in the methodology section.

Task structure is considered by Fiedler to be the second most important situational variable. The leader's control is increased as task structure is increased (Fiedler & Chemers, 1974). The extent to which the task is structured or unstructured is determined by four components: decision verifiability; goal clarity; goal path multiplicity; and solution specificity, for example, a structured task is characterized by high decision verifiability, high goal clarity, path simplicity, and high solution specificity (Fiedler, 1967). Task structure is determined by having the leader complete the Task Structure Rating Scale (TS). The scale consists of 10 items covering the four components of task structure. In this context Fiedler views task structure in terms of the task the group is required to perform. To the extent that the group successfully accomplishes the task, the leader is said to be effective. In this study this refers to the degree to which teachers completed the task of successfully achieving the school effectiveness components.

Position power is viewed by Fiedler as the least important of the situational variables. This variable is the degree to which the position itself permits the leader to gain compliance from group members. This definition closely aligns with French and Raven's (1958) concepts of legitimate power and reward and punishment power (Fiedler, 1967). The contention is that the more position power the leader has, the greater will be his or her control and influence over group members. It was Fiedler's (1967) belief that it was rarely necessary to rate position power, since almost all leaders have high position power. However, to accomplish the task, a Position Power (PP) questionnaire was developed by Fiedler. This aspect has always been somewhat controversial in the literature as it relates to schools (Crehan, 1984). Whether principals are considered to have high position power in schools is consistently a topic of debate.

To complete the situational favorability dimension, Fiedler (1967) a priori dichotomized the situational variables into good and poor leader-member relations, structured and unstructured tasks, and strong and weak position power. These were combined to yield eight different categories called *octants*. The eight octants were then arranged by Fiedler in an order that represented the amount of control and influence provided for the leader by each situation, or the degree of favorableness for the leader, as represented in Figure 1.

During Fiedler's (1967) earliest work in the area of psychology, he was interested in designing an instrument to measure interpersonal relations. Fiedler noted that there was a positive correlation between the competency of leader and the perceived similarity between leaders and followers. In his research, respondents were asked to describe their least preferred co-worker.

Octant	I	II	III	IV	V	VI	VII	VIII
Group atmosphere	good	good	good	good	poor	poor	poor	poor
Task structure	high	high	low	low	high	high	low	low
Position power	strong	weak	strong	weak	strong	weak	strong	weak

Figure 1. Favorableness for the leader concept (Fiedler, 1967, p. 176).

It appeared that how people view their co-workers gave an indication of their fundamental views of working in an organization. With respect to leaders it appeared that their views of co-workers was related to their underlying approach to administration, that is, if leaders viewed their co-workers in a positive way, they tended to be humanistic administrators. Fiedler's research led to the development of the Least Preferred Co-worker Scale (LPC). To complete the instrument, each respondent is asked to think of all the people with whom he or she has ever worked, then to focus on the one person with whom he or she found it the most difficult to cooperate. This person is known as the least preferred co-worker. This person can be someone the respondent is currently working with or has worked with any time in the past. To describe that person, the respondent completes a 16-item semantic differential scale. An example of the LPC scale is as follows:

Pleasant	8	7	6	5	4	3	2	1	Unpleasant
Friendly	8	7	6	5	4	3	2	1	Unfriendly
Open	8	7	6	5	4	3	2	1	Guarded

Respondents indicate where on this range of eight points they would place their target person. Thus a score on the 16 items could range from a low of 16 to a high of 128. According to Fiedler (1977) a score below 58 indicates a low LPC leader, that is, an individual who describes his or her co-worker in a negative fashion would be described as a task-oriented leader. A score above 63 indicates a high LPC leader. The person in this category describes his or her co-worker in a positive fashion and would be described as a relation-oriented leader.

Throughout the Fiedler literature LPC scores have been interpreted as an index of relationship-versus-task orientation; the high LPC individual is depicted as relationship-oriented, while the low LPC individual is believed to be task-oriented (Fiedler, 1973). The implication that leaders' regard for co-workers reflects their fundamental views of administration has not received unanimous support (Crehan, 1984). Although some studies support the interpretation of LPC scores as an index of relationship-versus-task orientation, others have expressed strong doubt. Fiedler attempts to reconcile these major differences by stating:

> Leaders tend to behave in a human-relations-centred manner in one situation but in a job-oriented manner in another. This means that we cannot define leadership style by leadership behavior. There are no generally considerate

> leaders, only those who are considerate in some situations and inconsiderate in others. (Fiedler & Chemers, 1974, p. 102)

A motivational interpretation of LPC was proposed by Fiedler (1977) to account for these fluctuations. This suggests that people have different goals but also their motivation to achieve them corresponds to the importance they attach to each goal. Therefore, Fiedler (1973) states that high LPC persons are motivated primarily to achieve successful interpersonal relations. Their secondary goals include self-enhancement, prominence, and esteem for others. The primary goal of low LPC persons is directed toward goal attainment. Their secondary goals focus on interpersonal relations, especially if such relations will contribute to task achievement. When the situation is stressful, the high LPC leaders will seek to accomplish the goals most important to them, the needs for affection, approval, and status, while the low LPC leaders will focus on their most important goals, the needs involving goal attainment. Under favorable conditions when leaders know that their primary goals will be achieved, high LPC leaders will focus on their task accomplishments and on seeking status and esteem from others, for example, their superordinates, while low LPC leaders will focus on being considerate and person-oriented.

As it applies to the school context, Fiedler's theory predicts that low LPC principals are more effective than high LPC principals when the situation is favorable or unfavorable, and that high LPC principals are more effective than low LPC principals when the situation is moderately favorable.

In spite of continued controversy, Fiedler's leadership theory has generated voluminous studies in education. Since the McNamara (1968) study, additional researchers (Faulkner, 1981; Garland, 1973; McKague, 1970; Thomson, 1972;) have investigated the relationships among the principal's leadership style and a variety of variables. These variables included such topics as level of bureaucratization, school achievement on standardized tests, organizational climate, and school supervisory practices. Crehan (1984), in a meta-analysis of the research, found over 400 studies using Fiedler's theory, either partially or completely. She was somewhat critical of studies in educational settings, citing methodological concerns. Her final comment was that it was not possible to draw conclusive evidence for either support or nonsupport for the theory. An analysis of the literature (Oakley, 1989) indicated that the Fiedler model has not been related to the variable of school effectiveness.

The present study was designed to test Fiedler's Contingency theory using school effectiveness as a measure of group effectiveness. The study could provide some possible modifications as proposed by Crehan (1986) or indicate that the Fiedler theory appears to be of dubious value in predicting school effectiveness. For this study, group effectiveness consisted of the following seven components as identified in the school effectiveness literature: strong leadership on the part of the principal; high expectations by the principal and staff for student achievement; a system for monitoring student progress; a positive, orderly climate; emphasis on basic skill achievement; clearly stated, well-defined goals; and good home-school relations. Schools designated as effective, using a variety of criteria, have been found to have these characteristics.

Hypotheses

1. There will be a negative correlation between LPC scores of elementary principals and school effectiveness when the situational favorability for the principal is high.
2. There will be a positive correlation between LPC scores of elementary principals and school effectiveness when the situational favorability for the principal is moderate.

Instrumentation

A measure of school effectiveness was required that contained items reflecting the seven components previously identified in the literature as characterizing an effective school. A review by Guzzetti (1983) identified several instruments in use in school districts; however, most had been designed for school improvement projects. Because the proposed study was descriptive in nature and few reliability and validity measures were available on existing instruments, it was decided to modify Bohac's (1985) School Effectiveness Questionnaire. Bohac had validated her instrument with a random sample of urban and rural schools in Canada in a similar population to that proposed for this study.

Bohac (1985) developed her questionnaire based on the conceptual framework developed by Steers (1976), as well as on criteria identified from the organizational and school effectiveness literature. Using a panel of judges, she reduced an initial list of over 100 criteria to 60. Following extensive reliability and validity measures, a 37-item questionnaire was developed. Because the Bohac instrument was heavily loaded on two dimensions for her specific study, a validity and reliability process was undertaken to modify the instrument for the proposed study. Additional items were added to the instrument from literature findings so that the first draft of the modified instrument contained 49 items.

A panel of judges was selected to assist in the validation process. The judges were asked to categorize each item into the appropriate seven components of school effectiveness. It was decided to include only those items on which six or more judges agreed on the classification. The judges were also asked to identify any ambiguous items. This process resulted in rewording a number of items and the elimination of 12 items. The revised questionnaire was administered to a group of teachers who were asked to complete the questionnaire in relation to their school. This procedure did not result in any changes to the questionnaire.

To ensure reliability of the instrument, it was administered to two classes of 24 teachers who were enrolled in graduate classes. Internal consistency reliability was determined using Cronbach's alpha coefficient. The overall reliability coefficient for the questionnaire was 0.95. In two of the seven components it was found that the alpha could be significantly increased by eliminating one item. Therefore, the final version of the questionnaire had 35 items with the breakdown and alpha coefficients shown in Table 1.

As previously discussed, the Least Preferred Co-worker Scale (LPC) is a 16-item semantic differential, Likert-type scale developed by Fiedler (1964). LPC scores have a high internal consistency, with split-half coefficients of over .90 reported. Gruenfeld, Rance, and Weissenberg (1969) reported a test-retest coefficient of .85 (n = 126) reported over a five-week period. Results in this

Table 1
Reliability Coefficients and Frequency of Items of the School Effectiveness Questionnaire

Components	*Reliability Coefficient*	*Frequency of Items*
Principal leadership	.84	7
High expectations	.58	3
Positive, orderly climate	.91	9
Emphasis on basic skills	.76	4
Frequent, systematic evaluation of students	.75	3
Clearly stated goals	.77	4
Good home-school relations	.83	5

study were relatively widely distributed, allowing us to consider scores of 64 and above as high LPC leaders, while scores of 57 and below were considered low LPC leaders.

The Group Atmosphere Scale (GA), consisting of a 10-item semantic differential scale, was developed by Fiedler (1967) who reported a split-half reliability of .90. Schools with scores of 65 and below were considered to have poor leader-member relations while scores of 66 and above were considered to have good leader-member relations.

The Principal Position Power questionnaire, consisting of 12 items, developed by Hunt (1967) and modified by Bobner (1982) was used. The average interrater agreement among judges rating position power was .95 (Bobner, 1982). Scores of six and above were considered to indicate strong position power while scores of five or below were considered to indicate weak position power (Hunt, 1967).

The Task Structure Rating Scale used in the study was developed by Fiedler and Chemers (1974). The questionnaire consists of 10 items based on the components outlined previously. Responses can range from 0 to 20 with a score of 6 below being unstructured, a score of 7 to 13 being moderately structured, and a score of 14 or above being structured. We were unable to obtain reliability statistics for the scale, since it had seldomly been used in educational organizations.

Sample

The data were collected from schools in 33 of the 34 school districts in the province of Newfoundland and Labrador. Principals were asked to complete the least preferred co-worker scale, the group atmosphere scale, the task structure questionnaire, and the position power questionnaire. This procedure corresponds with the suggestion made by Fiedler and Chemers (1974). Teachers' data were collapsed into a school mean to determine a school effectiveness score.

The questionnaires were mailed to 281 schools, with 184 schools responding. Due to incomplete data, eight schools had to be discarded, leaving a sample of 176 schools and principals. School effectiveness instruments were returned by 899 teachers.

Table 2
Descriptive Statistics for Leadership Style (LPC), Group Atmosphere, Task Structure, Position Power, and School Effectiveness

Variable	*Mean*	*Standard Deviation*	*Median*	*Mode*	*Range*
Leadership style	60.90	18.54	62	62	16-125
Group atmosphere	65.49	8.28	66	70	37-80
Task structure	10.39	3.31	11	11	0-17
Position power	8.69	1.57	9	10	4-12
School effectiveness	163.78	22.41	166	175	64-210

The hypotheses were analyzed using the Spearman rank order correlation formula.

Because there were only four schools where the principal rated the position power as weak, and these were distributed over three octants, these scores were eliminated from the analysis. As a result of this classification there were no schools in octants II, IV, VI, and VIII. A total of 117 principals rated task structure as moderate and therefore could not be used in the analysis. The final allocation placed 26 schools in octant I, 5 schools in octant III, 6 schools in octant V and 18 schools in octant VII. Therefore, 31 schools rated as having high favorability, and 24 schools rated as having moderate favorability.

Hypotheses 1 and 2 represent essentially the main predictions of Fiedler's Contingency theory with respect to the school situation. Fiedler's theory states that low LPC principals will have higher group effectiveness than high LPC principals when there is high situational favorability, whereas high LPC principals will have higher group effectiveness than low LPC principals when the situational favorability is moderate. Neither of these hypotheses had a correlation that was statistically significant. From an inferential statistics perspective, no confidence can be placed in these two correlations. The correlation obtained for hypothesis 1 (.19, n 31) was directionally incongruent with that predicted by the Contingency Theory. The theory predicts negative correlations for octants I (good GA, high TS, and strong PP) and III (good GA, low TS, and strong PP). This result appears to confirm the suggestion made by Crehan (1984) that studies using the theory in schools have a low incidence of conformity. Because we consider our dependent variable, school effectiveness, to be reliably measured, one must question the use of Fiedler's theory in a modern educational context.

The correlation obtained for hypothesis 2 is .29 (n = 24), which is directionally congruent with the predictions of the theory. The theory predicts positive correlations of .42 and .05 for octants V (poor GA, high TS, and strong PP) and VII (poor GA, low TS, and strong PP). Our result is consistent in significance with other results reported in the literature involving school studies. McNamara (1968) reported a positive correlation of .31 for octant V, while Bobner (1982) found directional congruence for both octants V and VII. The magnitude of the correlations was not reported. It is difficult to place too much credibility on the findings of directionality. Research on effective schools

suggests that there may be variables operating at both the school level and the classroom level affecting the level of effectiveness (Crawford & Raia, 1988). Since the contingency theory posited by Fiedler is measuring variables at the school level, the instruments may be only partially obtaining the full affects of school and classroom practices.

Discussion

The results of this study challenge the main postulate of Fiedler's (1967) theory that group performance is contingent on the appropriate matching of leadership style and the degree of favorableness of the group situations for the leader. Given the moderate size of the sample, the rigor of the research methods, and strict adherence to the theoretical construct, it does not appear that the model is appropriate for today's school situations. Perhaps in the past there were school variables that related in some way to the model; however, the concept of school effectiveness appears to behave otherwise. In the following paragraphs an attempt is made to suggest some apparent discrepancies in the model and propose some required modifications to accommodate school effectiveness variables.

It is difficult to argue that a situational aspect is not relevant to school leadership. Current literature in program implementation, school culture, school effectiveness, and other variables all suggest that a situational component exists in school leadership.

Fiedler's major instrument for assessing leadership style, the Least Preferred Co-worker Scale (LPC) has been the subject of continued controversy since its development. Although the level and distribution of scores were appropriate, it is doubtful that the instrument is measuring leadership style. We believe that there are excellent leaders in effective schools who are both task-oriented and relation-oriented. It is doubtful that their style would relate to the situational variables proposed in the Fiedler model.

Although the major concern in this study was related to the LPC component, a number of issues regarding the situational variables should be addressed. According to the theory, Group Atmosphere has the most impact on situational favorability. The scores in this study were based on principals' reports and ranged from 31 to a perfect score of 80. This spread of results causes some concern. One must be skeptical when a principal rates his or her group atmosphere as the highest possible. An issue that has received some attention in the literature and that deserves a closer look is the possibility that the perception differs between leaders and members. Posthuma (1970) reported that there was no significant difference between leaders' and group members' perceptions of leader-member relations in experimental groups. In an early report, Fiedler (1967) found that the relationship between the leaders' and members' perceptions of group atmosphere was low. In the meta-analysis conducted by Strube and Garcia (1981), they reported that there was no regular pattern for assessing group atmosphere. On some occasions it was measured by the leader, while on other occasions it was measured by the members. Other leadership models (e.g., Likert, 1967) are very definite in this aspect and state that member relations should be assessed by subordinate perception.

Our criticism of the assessment of group atmosphere goes beyond the question of whose perception is presented. The major concern expressed in this study is that there is more than leader-member relations involved in group atmosphere on a school situation. It is our contention that this variable must be broadened to include other aspects, such as peer relationships, student impacts, and other cultural components. The group atmosphere variable has the potential to be one of the major situational variables. As it is currently defined and measured in Fiedler's model, it is not sophisticated enough for today's schools.

The structure of the task the group is required to perform is the second most important variable. Respondents are required to give their perceptions of leader influence over group performance, in the sense that a structured task enables the leader to impose organizational controls on the work process. As discussed earlier in the article, Fiedler has identified two kinds of task groups: those in which the task requires interdependent action by the group members, known as interacting groups, and those in which the group members work independently, known as coacting groups. According to Fiedler, the teaching process falls into the coacting group category. This study casts serious doubt on the applicability of this assumption in relation to the school situation. Current literature on school effectiveness (Mortimore et al., 1986) suggests that effective schools result when a highly motivated, highly integrated group of professionals work interdependently to accomplish organizational goals. If this is the case, there is no way that the school could be classified as made up of coacting groups. Crehan (1984) alludes to this discrepancy in her meta-analysis when she suggests that teaching may be both a structured and unstructured task depending on the nature not only of what is being taught, but also the pupil to whom it is being taught. Given this point of view, it is very difficult to designate the task of the school as structured or unstructured. The wide range of responses suggested that principals were somewhat uncertain how to classify the task. Principals may have perceived task structure based on the type of teaching approach used primarily in the school, their view of educational goals, or other perceptions. The wide range of results obtained suggests a variety of interpretations influenced the responses. More than 60% of the respondents placed their schools midway between structured and unstructured, a situation that Fiedler and Chemers (1974) called moderately structured. It is questionable whether a great deal of confidence can be placed in the credibility of this situational variable as it is currently defined and assessed.

The position power of the principal is viewed by Fiedler as the least important of the situational variables. This variable relates to the degree to which the position itself permits the principal to gain compliance from the staff. The majority of the principals in this study perceived their position power to be high; in fact only four principals rated their position power as low. In spite of the fact that this finding agrees with others in the literature (McKague, 1970; Mellor, 1974), one must question the designation when it is given by the principal. There is a possibility that others in the school or school system might not view the principal as having a lot of power. As a self-assessment, principals may want to give the impression that they occupy a seat of

power. Other views in the literature (Fullan, 1982) would suggest that much of the principal's power has been eroded due to collective bargaining and other sociological factors. In conclusion, although principal position power is only a minor variable in determining situational favorability, it is another instance where Fiedler's Contingency theory needs modification to fit today's school situations.

Conclusion

A meta-analysis conducted by Crehan (1984) found some support for the Contingency model in organizational settings other than schools. At that time she suggested that the applicability of the model to the school setting was tentative.

Results of this study tend to agree with Crehan's (1986) comments. It appeared that the school effectiveness dimension was a reliable variable to demonstrate the applicability of Fiedler's model to the school situation. A strong situational dimension appears to be implied in the current views of school effectiveness. As one of the major contingency approaches in the literature, it was felt that a test of the theory was warranted. However, the results suggest that major modification of the theory is necessary before it can be rigorously applied to the school situation.

Initially, there are major concerns with the concept of least preferred coworker (LPC) as it applies to schools. Leaders in todays schools must be able to demonstrate both task and relationship behavior. Contingency models must reflect this finding. Refinement and modification of the situational variables must take place before the Fiedler model can be applied successfully to school situations. In this context, it is suggested that Group Atmosphere be measured by subordinates in educational settings.

The definition of task structure must be clarified. The majority of the respondents in this study depicted the task structure of schools as moderate. It is possible that the data in the middle reflect an organization containing both structured and unstructured tasks.

Similarly, clarification is required for the position power variable. Literature on the concept of empowerment in education would indicate that principals have relatively low position power. Perhaps this variable should also be rated by subordinate perception. Any attempts at matching schools with the leadership style of principals based on the use of the Fiedler theory must be considered futile. Major modifications are required before further research can be justified.

References

Austin, G. (1978). *Process evaluation: A comprehensive study of outliers.* Maryland State Department of Education and Center for Educational Research and Development, University of Maryland. (ERIC Document Reproduction Service No. ED 160 644)

Bobner, R.F. (1982). *Situation favorableness as a variable of leadership effectiveness in selected secondary schools in Ohio.* Unpublished doctoral dissertation, University of Akron.

Bohac, V.J. (1985). *School effectiveness criteria.* Unpublished doctoral dissertation, University of Alberta.

Brookover, W.B., & Lezotte, L.W. (1979). *Changes in school characteristics coincident with changes in student achievement.* East Lansing: Institute for Research on Teaching, Michigan State University. (ERIC Document Reproduction Service No. ED 181 005)

Coleman, J.S., Campbell, E.G., Hobson, C.J., McPartland, J.M., Mood, A.M., Weinfeld, F.D., & York, R.L. (1966). *Equality of educational opportunity*. Washington, DC: Government Printing Office.

Collard, J.L. (1984). School effectiveness: A re-examination of the research evidence. *Journal of Educational Administration, 22*(2), 146-155.

Crawford, J., & Raia, F. (1988). *Results from correlating school-level educational indicators and process measures with achievement*. Paper presented at American Educational Research Association, New Orleans.

Crehan, E.P. (1984). *A meta-analysis of Fiedler's contingency model of leadership effectiveness*. Unpublished doctoral dissertation, University of British Columbia.

Crehan, E.P. (1986). A meta-analysis of Fiedler's contingency model of leadership effectiveness: Some selected findings. *CSSE Occasional Paper Series*. Ottawa: Canadian Society for Study of Education.

Faulkner, R.T.B. (1981). *Least preferred co-worker scores of the principal and degree of bureaucratization in open and traditional schools*. Unpublished doctoral dissertation, University of Ottawa.

Fiedler, F.E. (1964). A contingency model of leadership effectiveness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 150-189). New York: Academic Press.

Fiedler, F.E. (1967). *A theory of leadership effectiveness*. New York: McGraw-Hill.

Fiedler, F.E. (1973). The contingency model—a reply to Ashour. *Organizational Behaviour and Human Performance, 9*(3), 356-368.

Fiedler, F.E. (1977). A rejoinder to Schriesheim and Kerr's premature obituary of the Contingency Model. In J.C. Hunt & L.L. Larson (Eds.), *Leadership: The cutting edge*. Carbondale, IL: Southern Illinois University Press.

Fiedler, F.E., & Chemers, M.M. (1974). *Improving leadership effectiveness: The leader match concept* (2nd ed.). New York: Wiley.

Fiedler, F.E., Chemers, M.M., & Mahar, L. (1976). *Improving leadership effectiveness: The leader match concept*. Toronto: Wiley.

Fishbein, M., Landy, E., & Hatch, G. (1969). A consideration of two assumptions underlying Fiedler's contingency model for the prediction of leadership effectiveness. *American Journal of Psychology, 82*, 457-473.

Fleishman, E.A. (1973). Twenty years of consideration and structure. In E.A. Fleishman & J.G. Hunt (Eds.), *Current developments in the study of leadership* (pp. 1-37). Carbondale, IL: Southern Illinois University Press.

French, J.R., & Raven, B.H. (1958). Legitimate power, coercive power and observability in social influence. *Sociometry, 21*, 83-97.

Fullan, M. (1982). *The meaning of educational change*. Toronto: OISE Press.

Garland, P. (1973). *The effect of principal-teacher interaction in secondary school environments: An empirical study*. Unpublished doctoral dissertation, University of Ottawa.

Geering, A. (1980). *The role of the school principal in comparative perspectives*. (ERIC Document Reproduction Services No. ED 199-930)

Goodman, S.M. (1959). *The assessment of school quality*. Albany, NY: State Department of New York.

Gruenfeld, L.W., Rance, E.E. & Weissenberg, P. (1969). The behaviour of task-oriented (low LPC) and social oriented (high LPC) leaders under several conditions of social support. *Journal of Social Psychology, 79*(1), 99-107.

Guzzetti, B.J. (1983). *Report on instruments for measuring school effectiveness*. Aurora, CO: Mid-continent Regional Educational Laboratory.

House, R.J., & Baetz, M.L. (1979). Leadership: Some empirical generalizations and new research directions. *Research in Organizational Behaviour, 1*, 341-423.

Hunt, J.G. (1967). Fiedler's leadership contingency model: An empirical test in three generations. *Organizational Behavior and Human Performance, 2*(3), 290-308.

Hunt, J.G. (1984). Organizational leadership: The contingency paradigm and its challenges. In B. Kellerman (Ed.), *Leadership: Multi-disciplinary perspectives*. Englewood Cliffs, NJ: Prentice-Hall.

Jencks, C.S., Smith, M., Ackland, H., Bane, M.J., Cohen, D., Gintis, H., Heyns, B., & Michelson, S. (1972). *Inequality: A reassessment of the effect of the family and schooling in America*. New York: Basic.

Leithwood, K.A., & Montgomery, D.J. (1981). *The role of the elementary school principal in program improvement: A progress report*. Paper presented at the Canadian Society for Studies in Education Annual Meeting, Halifax, NS.

Lewin, A.Y., & Minton, J.W. (1986). Determining organizational effectiveness: Another look, and an agenda for research. *Journal of Management Science, 32*(5), 514-538.

Likert, R. (1967). *New patterns of management*. New York: McGraw-Hill.

Lipham, J. (1977). *The administrator's role in educational linkage*. Columbus, OH: UCLA. (ERIC Document Reproduction Service No. ED 141 900)

McKague, T.R. (1970). LPC—A new perspective on leadership. *Educational Administration Quarterly, 6*(3), 1-14.

McNamara, V.D. (1967). *A descriptive-analytic study of directive-permissive variation in the supervisory behavior of elementary school of principals*. Unpublished master's thesis, University of Alberta.

Mellor, K.P. (1974). *An investigation of Fiedler's Contingency Theory of leadership effectiveness as it applies to elementary school principals in Rhode Island*. Unpublished doctoral dissertation, University of Connecticut.

Mollenkopf, W.G., & Melville, S.A. (1956). *A study of secondary school characteristics as related to test scores* (Research Bulletin 56-6). Princeton, NJ: Educational Testing Service.

Mortimore, P., Sammons, P., Stoll, H., Lewis, D., & Ecob, R. (1986). *The junior school project*. ILEA Research and Statistics Branch, Addington Street Annex, County Hall, London, SE1 7UY.

Nebeker, D.M. (1975). Situational favorability and environmental uncertainty: An integrative study. *Administration Science Quarterly, 20*, 281-294.

Oakley, W.F. (1989). *A study of the relationship between the leadership style of the principal and school effectiveness under differing conditions of situational favorability*. Unpublished doctoral dissertation, University of Ottawa.

Posthuma, A.B. (1970). *Normative data on the Least Preferred Co-worker scale (LPC) and the Group Atmosphere scale (GA)* (Technical Report 70-8). Seattle, WA: Organizational Research, Department of Psychology, University of Washington.

Purkey, S.C., & Smith, M.S. (1982). Too soon to cheer? Synthesis of research on effective schools. *Educational Leadership, 40*(3), 64-69.

Purkey, S.C., & Smith, M.S. (1983). Effective schools: A review. *Elementary School Journal, 83*(4), 427-452.

Rowan, B., Bossert, S.T., & Dwyer, D.C. (1983). Research on effective schools: A cautionary note. *Educational Researcher, 12*(4), 24-31.

Rutter, M., Maughan, B., Mortimore, P., Quston, J., & Smith, A. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. Cambridge, MA: Harvard University Press.

Sackney, L.E. (1986). Practical strategies for improving school effectiveness. *Canadian School Executive, 6*(4), 21-26.

Schermerhorn, J.R., Hunt, J.G., & Osborn, R.N. (1982). *Managing organizational behavior*. New York: Wiley.

Shoemaker, J., & Fraser, H.W. (1981). What principals can do: Some implications from studies of effective schooling. *Phi Delta Kappan, 63*(3), 178-182.

Steers, R.M. (1976). When is an organization effective? A process approach to understanding effectiveness. *Organizational Dynamics, 5*, 50-63.

Stogdill, R.M. (1948). Personal factors associated with leadership: A survey of the literature. *Journal of Psychology, 25*, 35-71.

Strube, M.J., & Garcia, J.E. (1981). A meta-analytic investigation of Fiedler's model of leadership effectiveness. *Psychological Bulletin, 90*(2), 307-321.

Thomson, J.G. (1972). *An empirical study of the relationship between Ontario secondary school principals' leadership effectiveness and helping relationship in teachers*. Unpublished doctoral dissertation, University of Ottawa.

Weber, G. (1971). *Inner-city children can be taught to read: Four successful schools*. Washington, DC: Council for Basic Education.

Wright, R., & Renihan, P. (1985). *The Saskatchewan principalship study: A review of the literature* (Report No. 127, #400-2222). Regina, SK: Saskatchewan School Trustees Association Research Centre.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 1, March 1992, 49-64*

B.E.J. Housego
University of British Columbia

Monitoring Student Teachers' Feelings of Preparedness to Teach, Personal Teaching Efficacy, and Teaching Efficacy in a New Secondary Teacher Education Program

Student teachers' feelings of preparedness to teach as measured on the PREP Scale (Housego, 1990b), personal teaching efficacy, and teaching efficacy as measured on the Gibson Scale of Teacher Efficacy (Gibson & Dembo, 1984) were monitored across the second offering of the revised UBC secondary teacher education program, 1988-1989, term by term, for the total group and subject specialization and gender subgroups. Total group and all subgroup PREP, personal teaching efficacy, and teaching efficacy mean scores increased significantly in first term. In second term total group mean PREP scores increased significantly but mean personal teaching efficacy and teaching efficacy scores did not. In third term neither mean PREP scores nor personal teaching efficacy nor teaching efficacy mean scores increased significantly. The findings for the four subject specialization subgroups were similar. All PREP Scale mean scores increased significantly in first term, two increased significantly in second term and none increased significantly in third term. Two subgroups recorded significant increases in personal teaching efficacy in first term; none were recorded beyond that point. No subject specialization subgroup recorded significant increases in teaching efficacy at any point; indeed, one subgroup recorded significant decreases during terms two and three. Both male and female PREP Scale mean scores increased in terms one and two but not three. Personal teaching efficacy mean scores increased significantly for both gender groups in first term, for females in second term (the practicum), and for neither group in third term. Neither subgroup recorded significant increases in teaching efficacy at any point. Implications for teacher education are discussed.

Les sentiments d'élèves-maîtres concernant leur état de préparation à l'enseignement mesurés par l'échelle PREP *(Housego, 1990b), ceux concernant leur efficacité personnelle à l'enseignement, et l'efficacité à l'enseignement d'après eux, mesurés par l'échelle* Teacher Efficacy *de Gibson (Gibson & Dembo, 1984) étaient suivis pendant le second déroulement (1988-1989) du programme revisé de formation de maîtres du cycle secondaire de l'Université de la Colombie Britannique. Cela était fait pendant chaque trimestre pour tout le groupe et pour les sous-groupes d'élèves formés à partir des matières de spécialisation et du sexe. Les moyennes des scores de l'échelle PREP, de l'efficacité personnelle à l'enseignement, et de l'efficacité à l'enseignement augmentèrent d'une manière significative au premier trimestre. Pendant le second trimestre, les moyennes des scores de l'échelle PREP pour tout le groupe augmentèrent d'une manière significative, mais celles de l'efficacité personnelle à l'enseignement et l'efficacité à l'enseignement n'augmentèrent pas. Pendant le troisième trimestre, ni les moyennes des scores de l'échelle PREP, ni celles de l'efficacité personnelle à l'enseignement et l'efficacité à l'enseignement n'augmentèrent d'une manière significative. Les résultats pour les quatre sous-groupes de matières de spécialisation étaient semblables. Toutes les moyennes des scores de l'échelle PREP augmentèrent d'une manière significative au premier trimestre, deux moyennes augmentèrent au second trimestre, et aucune d'elles n'augmenta au troisième trimestre.*

Billie Housego is an associate professor in the Department of Educational Psychology and Special Education, and she is associated with the Centre for the Study of Teacher Education.

Deux sous-groupes ont connu des augmentations significatives dans les moyennes de leurs scores d'efficacité personnelle à l'enseignement au cours du premier trimestre; aucune autre augmentation ne fut constatée au-delà du premier trimestre. Aucun sous-groupe de matières de spécialisation n'a connu des augmentations significatives dans l'efficacité à l'enseignement à aucun moment; d'ailleurs, pour un des sous-groupes on enregistra des diminutions significatives pendant les deuxième et troisième trimestres. Les moyennes des scores de l'échelle PREP *pour les garçons aussi bien que pour les filles augmentèrent au cours du premier trimestre et du second, mais pas au troisième trimestre. Les moyennes des scores de l'efficacité personnelle à l'enseignement augmentèrent d'une manière significative aussi bien pour les garçons que pour les filles pendant le premier trimestre; pour les filles pendant le second trimestre (le stage d'enseignement pratique), et pour aucun groupe pendant le troisième trimestre. A aucun moment, pour l'un ni l'autre sous-groupe on n'enregistra une augmentation des scores d'efficacité à l'enseignement. Les implications de cette recherche pour la formation des enseignants sont présentées.*

Self efficacy, an important personal characteristic or disposition, is described by Bandura (1986, p. 391) as the "judgment of one's capacity to accomplish a certain level of performance." It seems to imply a degree of preparedness by virtue of training, experience, or talent. It functions in a reciprocal relation to a specific behavior and the environment in which that behavior occurs (Bandura, 1978). Theoretically, change in a personal factor, for example, self efficacy regarding a teaching behavior such as gaining attention or explaining a particular concept, would affect the teaching behavior, which in turn would alter the educational environment. Each component plays a role in shaping the other two. Bandura (1977, p. 191) proposes that self efficacy "determines whether coping behavior will be initiated, how much effort will be expended and how long it will persist in the face of aversive experiences." He also distinguishes efficacy expectations from outcome expectations. Through the latter, one understands that a particular course of action will produce certain outcomes, whether or not one can perform those actions oneself. One is likely to act on an outcome expectation when one expects to be successful performing the behavior required to attain it, that is, one has the necessary efficacy expectation.

Two studies funded by the Rand Corporation (Armor et al., 1976; Berman, McLaughlin, Bass, Pauly, & Zellman, 1977) distinguished between outcome and efficacy expectations regarding teaching. They measured efficacy as the total score on two Likert-style items:

1. When it comes right down to it, a teacher really can't do much because most of a student's motivation and performance depends on his or her home environment.
2. If I really try hard, I can get through to even the most difficult and unmotivated students.

These two items have been found to be not significantly correlated (Ashton, Olejnick, Crocker & McAuliffe, 1982; Poole & Okeafor, 1989; Tracz and Gibson, 1986; Woolfolk, Rossoff & Hoy, 1989). The distinction between the outcome expectation of the efficacy of teaching (Rand 1 efficacy) and the teacher's expectation of personal competence (Rand 2 efficacy) is therefore supported. Ashton and Webb (1986, p. 140), refer to outcome expectations as teaching efficacy (a "multi-dimensional situation specific construct" describing what it is possible to accomplish) and efficacy expectations as personal teaching efficacy (what the individual teacher, given his or her characteristics and the

current environment, feels he or she can accomplish or is prepared to accomplish). This article uses their terminology.

Researchers have established relationships between pupil achievement and both teaching efficacy and personal teaching efficacy. Beyond the Rand Corporation studies (Armor et al., 1976; Berman et al., 1977), in which teacher efficacy beliefs were shown to be related to pupil reading and mathematics achievement, Tracz & Gibson (1986, p. 3) reported positive correlations between personal teaching efficacy and pupil reading achievement and between teaching efficacy and pupil mathematics achievement. Also investigating the relationship between teacher efficacy and student learning responses, Midgely, Feldlaufer, and Eccles (1988) found that pupils in transition from elementary to junior high school who moved from the classroom of a high efficacy teacher to the classroom of a low efficacy teacher demonstrated lower perceived performance expectations, lower perceived performance, and higher expectations of difficulty in mathematics than pupils who moved from one low efficacy teacher to another. Similarly, Anderson, Greene, and Loewen (1987) identified personal teaching efficacy at the beginning of the school term as a significant factor in the achievement of grade 3 pupils and in the development of self efficacy in both grade 3 and grade 6 pupils. Ashton and Webb (1986, p. 139) concluded that their findings "strongly support the hypothesis that teachers' sense of efficacy is related to student achievement."

Relationships have also been found between teaching efficacy or personal teaching efficacy and holding high expectations for pupils, maintaining on-task behavior, concentrating on academic instruction, and demonstrating "withitness" (Ashton, Webb, & Doda, 1983); using less small-group instruction (Gibson & Dembo, 1984; Tracz & Gibson, 1986); consulting parents and involving them in their children's education (Hoover-Dempsey, Bassler, & Brissie, 1987); relating to pupils in a more trusting and accepting manner (Ashton, 1984); and adopting a more humanitarian approach to pupil control (Woolfolk & Hoy, 1988; Woolfolk et al., 1989). Teachers with higher efficacy scores were rated more successful by superintendents (Trentham, Silvern, & Brogdon, 1985); found teaching less stressful (Parkay, Olejnik, & Proller, 1988; Greenwood, Olejnik, & Parkay, 1990); were less likely to leave the profession (Glickman & Tamashiro, 1982); and were less likely to label problem children and refer them to others for teaching (Meijer & Foster, 1988; Miller, 1988).

Since "high efficacy teachers tend to be teachers who exhibit the effective teaching behaviors found in the process-product research literature" (Parkay, et al., 1988, p. 2), it is to be expected that enhancing the development of teachers' sense of efficacy is becoming a common goal of staff development programs (Miller, 1988; Poole & Okeafor, 1989). It is a variable to which teacher educators should also attend (Ashton, 1984).

Few studies of the teaching efficacy and personal efficacy of student teachers have been done. Evans and Tribble (1986), studied perceived teaching problems, commitment to teaching, and self efficacy in elementary and secondary post-baccalaureate student teachers using the earlier 30-item version of the Gibson Teacher Efficacy Scale (Gibson & Dembo, 1984). They located gender and grade level effects. Female student teachers recorded higher efficacy scores than males, and elementary student teachers recorded higher

scores than secondary ones. Similar differences have been reported among practicing teachers (Cavers, 1988; Greenwood, Olejnik, & Parkay, 1990). Such findings raise questions about the development of feelings of preparedness for teaching, teaching efficacy, and personal teaching efficacy during the course of preservice teacher education.

The University of British Columbia teacher education programs were recently revised. Important modifications included a lecture type, generic teaching skills course with small group laboratories, a developmental practicum sequence culminating in a 13-week extended practicum, laboratory components in some other courses, and a communications course. The 12-month, three-term, post-degree secondary program begins with a term including the generic teaching skills course, a course analyzing educational issues, an educational psychology and adolescent development course, and subject specific methods courses. Second term includes the extended practicum preceded by a two-week communications course. The final term includes an educational psychology course, a foundations course, a school organization course, an English education course, and professional and academic electives. Program revision was planned by the faculty and subjected to the scrutiny of the wider university community, practicing teachers, and administrators and their professional organizations.

One might anticipate that appropriate program changes would enhance student teachers' feelings of preparedness to teach, their personal teaching efficacy (what teachers, given their individual characteristics and the current environment, feel they can accomplish or are prepared to accomplish), and their teaching efficacy (the expectation that it *is* possible to ensure learning gains through teaching), these being, perhaps, very basic indicators of the successful functioning of a program in which students should come to feel increasingly well prepared to teach and maintain high expectations regarding the goals to be attained through teaching.

This study monitors, for the total group and subgroups based on subject specialization and gender, student teachers' feelings of preparedness to teach, teaching efficacy and personal teaching efficacy from September 1988 to August 1989, the second offering of the revised secondary teacher education program at UBC. The PREP Scale (Housego, 1990b) was used to measure feelings of preparedness to teach and the Gibson Scale of Teacher Efficacy (Gibson & Dembo, 1984) to measure the other two variables. Both scales are described in the next section of the article. The following research questions were asked:

1. Did secondary student teachers' feelings of preparedness to teach increase significantly during each term of the program?
2. Did each secondary subject specialization and gender subgroup of secondary student teachers increase significantly in feeling prepared to teach in each term?
3. Did any subject specialization or gender subgroup of secondary student teachers feel significantly more prepared to teach than any other at the outset of the program or at the conclusion of any term?
4. Did teaching efficacy and personal teaching efficacy mean scores of secondary student teachers increase significantly in each term?

5. Did the teaching efficacy and personal teaching efficacy mean score of each subject specialization and gender subgroup of secondary student teachers increase significantly each term?
6. Did any subject specializaion or gender subgroup of secondary student teachers record a significantly higher mean teaching efficacy or personal teaching efficacy score at the outset of the program or at the conclusion of any term?

Method

Participants

Secondary student teachers, all of whom held degrees and enrolled to prepare themselves to teach one or two secondary school subjects, participated in this study. They were grouped in four subject specialization subgroups, paralleling those in a previous study (Housego, 1990a) as well as the departmental structure of the Faculty (see Table 1). Feelings of preparedness to teach, teaching efficacy, and personal teaching efficacy scores were obtained from the total group and subgroups at program entrance (baseline) and at the conclusion of each of the three terms. The actual and possible participation numbers and percentages for four data collections—September 1988; December 1988; May 1989; and August 1989—were as follows: 177 of 206 (86%), 129 of 202 (64%), 121 of 185 (65%), and 123 of 177 (69%). Table 1 displays the number of participants in each subject specialization and gender subgroup at each data collection point.

The Scales

The Student Teachers' Feelings of Preparedness to Teach Scale (PREP Scale) was designed to measure the degree to which student teachers feel prepared to perform a set of tasks outlined by an 11-member interdepartmental Faculty committee as core in a teacher education program, applicable across grade levels and subject matter fields, and worthy of inclusion in the first basic course of a professional education program. The PREP Scale, which has under-

Table 1
Number of Participants by Subject Specialization and Gender: Baseline (B) through Term Three

	Term			
Subject Specialization	*B*	*1*	*2*	*3*
Math, Science, Business. Ed. Home Ec. & Industrial Ed.	91	67	66	66
English, Social Studies, and other languages	56	32	29	34
Art, Music & Theatre	17	15	14	10
Physical Education	13	15	12	13
TOTAL	177	129	121	123
Gender				
Female	72	60	52	46
Male	105	69	69	77
TOTAL	177	129	121	123

gone one revision in order to align it with current student teaching evaluation forms, now includes 50 instead of the original 43 items, each stated so as to complete the sentence, "I feel prepared to ..." In revision, the five-point scale anchored on one end by "very well prepared" and on the other by "very poorly prepared" was replaced with a seven-point scale anchored on one end by "almost completely prepared" and on the other by "almost completely unprepared." In repeated administrations the revised PREP Scale, like the original, has been found to be very reliable (.95 based on Hoyt's coefficient, an index of item homogeneity) and valid for the purpose of this study (Housego, 1990a).

The Gibson Teacher Efficacy Scale (Gibson & Dembo, 1984) was also used. Nine six-point Likert-style items measure personal teaching efficacy, a teacher's belief that *he* or *she* has the skills and abilities necessary to bring about student learning and can in fact do so (Dembo & Gibson, 1985, pp. 174-175). The other seven items measure teaching efficacy, a somewhat more basic belief in the efficacy of teaching itself, the extent to which one believes that the effects of negative home and family background influences can be overcome through teaching, regardless of whether or not one is capable of doing it oneself.

Scoring and Analysis

Confidence interval data were obtained on the responses of the total group and subgroups of subjects on the PREP Scale and the Gibson Teacher Efficacy Scales four times during the program. In order to assess total or subgroup increases in feelings of preparedness, teaching efficacy, and personal teaching efficacy during each term, the significance of differences between successive mean scores was ascertained using a one-factor fixed effects analysis of variance design. The Newman-Keuls procedure was used to analyze post-hoc differences.

Report and Discussion of Findings

The findings of this study are presented in four subsections: (a) Total Group Mean PREP Scale Scores (b) Subject Specialization and Gender Subgroup Mean PREP Scale Scores (c) Total Group Mean Teaching Efficacy and Personal Teaching Efficacy Scores (d) Subject Specialization and Subgroup Mean Teaching Efficacy and Personal Teaching Efficacy Scores. Limitations of the study are discussed in a fifth subsection.

Total Group Mean PREP Scale Scores

The PREP Scale mean scores of the total group of secondary student teachers increased significantly (F=210.12, df=3/546, α =.05) during the first and the second, but not the third term, as shown in Table 2.

Like other preservice teachers (Book, Byers, & Freeman, 1983; Housego, 1990a; Lanier & Little, 1986; Weinstein, 1988), secondary UBC student teachers believed they knew a good deal about teaching before they began professional training. At program entry, they scored, on the average, 4 or more points on the 7-point preparedness scale on 24 of 50 items, and 3.5 or more points on 42 of 50 items. Relatively high ratings of preparedness for activities with which individuals have limited familiarity are to be expected when they base their ratings on vicarious as opposed to enactive (real life) experiences (Bandura, 1986, p. 223). Many students enter the Faculty having had experience working

Table 2
Total Group Mean PREP Scale Scores: Baseline (B) through Term Three

	Term			
	B	*1*	*2*	*3*
PREP Scale Mean Score	200.27	270.57*	294.83*	293.92

*Significantly different from the previous term score at the α=.05 level.

with children, and all, of course, were pupils and can readily see themselves in a teacher's role (Lortie, 1975; Weinstein, 1988, p. 33).

The absence of a significant increase in the average rating of feeling prepared to teach in term three, the final four months of the 12-month program, invites speculation. Term three follows the extended practicum, a very positive experience for many students who are able to demonstrate their teaching skills and affirm their vocational choices. Some resent having to return to do further coursework when competence "on the job" has been proven, and there is so little time and energy to prepare for the commencement of a first teaching assignment in September. Third term coursework includes: school organization in its social context; an educational studies elective—social foundations of education, educational philosophy, history, sociology or anthropology; educational psychology (a further study of learning, measurement and teaching); an English education course; and two or three electives chosen to supplement the student's subject matter specialization or professional interests. There may be an element of mismatch between this coursework, which is focused more on the context and content of teaching and less on teaching itself, and student concerns, which generally relate to what they perceive as pressing problems and more direct indicators of classroom success, such as pupil involvement (Harootunian & Yarger, 1981). Miklos and Greene (1987), studying Alberta teacher education programs, found foundations, educational administration, and educational psychology courses respectively the least, second least, and third least valued courses for teachers with two or three years experience. Kelly (1989) reported similar findings regarding teachers' reactions to Saskatchewan teacher education programs. Overall the final and less practical term probably seems something of a letdown for students; moreover, it lacks a sense of closure since students must pursue additional coursework to earn a Bachelor of Education degree.

Subject Specialization and Gender Subgroup Mean PREP Scale Scores

As revealed in confidence interval data, and as shown in Table 3, no subject specialization subgroup of secondary student teachers felt more prepared to teach than any other at the baseline or at the end of any term of the program.

There were significant PREP Scale mean score increases (F=42.33, df=15/554, α =.05) for every subject specialization subgroup in first term, as also shown in Table 3, and for two subgroups, including approximately 75% of the students, in second term. In third term, no subject specialization subgroup mean PREP Scale score increased significantly, a repetition of the total group finding.

Table 3
Subgroup Mean PREP Scale Scores by Subject Specialization and Gender: Baseline (B) through Term Three

Subject Specialization	B	Term 1	2	3
Math, Sci., Bus. Ed. Home Ec. & Ind. Ed.	197.54	269.61*	288.70*	288.39
English, Soc. St. and other languages	200.18	265.19*	299.31*	297.49
Art, Music & Theatre	209.24	275.67*	307.43	301.30
Physical Education	208.28	281.20*	303.00	290.08
Gender				
Female	193.72	273.23*	300.35*	298.52
Male	204.76	268.25*	290.67*	291.17

*Significantly different from the previous term score for the same subgroup at the α=.05 level.
+Significantly different at the α=.05 level from a corresponding subgroup mean score recorded at the same time.

Similarly, as revealed in confidence interval data and indicated on Table 3, no gender subgroup mean PREP Scale score differed significantly from that of the other gender at the baseline or at the end of any term of the program.

Both gender subgroup mean PREP Scale scores increased significantly (F=91.56, df=7/542, α=.05) in first and in second, although not in third, term.

Total Group Mean Teaching Efficacy and Personal Teaching Efficacy Scores

The Gibson Teacher Efficacy Scale provides additional access to student teachers' self-estimates of their "personal power and resources to affect pupils' classroom behavior in desirable ways" (Evans & Tribble, 1986). One component of the scale measures teaching efficacy, one's belief in teaching as a means of overcoming the effects of external factors which may impede student learning—home environment, family background, and parental influence. A second is the expectation that one is able to accomplish this oneself. The scale's author cautions that the two components should be considered separately in order to avoid "wash-out effects" (Personal communication, S. Gibson, January 24, 1990). Table 4 displays mean teaching efficacy and mean personal teaching efficacy scores at the baseline of the program and at the end of each term.

There are significant changes in the mean scores of both teaching efficacy (F=6.49, df=3/546, α=.05), which declined by the end of first term, and per-

Table 4
Total Group Mean Teaching Efficacy and Personal Teaching Efficacy Scores: Baseline (B) through Term Three

	B	Term 1	2	3
Teaching Efficacy	26.37	24.89*	23.90	24.29
Personal Teaching Efficacy	35.71	39.39*	40.50	40.20

*Significantly different at the α=.05 level from the previous term score.

sonal teaching efficacy (F=24.28, df=3/546, α=.05), which increased by the end of first term. Trend analysis of teaching efficacy scores revealed a declining trend. Interestingly, secondary student teachers indicated both a declining belief in teaching and an increasing confidence in their own abilities to teach within three months of program entry and before experiencing the extended practicum. Those who decry the effects of school experience on student teachers (Dewey, 1965; Lanier & Little, 1986; Zeichner & Tabachnick, 1981) would find no support in the finding that attitudes toward teaching appear to shift negatively in advance of school experience. Perhaps students quickly come to the realization that schools cannot do all they are expected to do.

Subject Specialization and Gender Subgroup Mean Teaching Efficacy and Personal Teaching Efficacy Scores

An overall test of subject specialization subgroup mean teaching efficacy scores and times (i.e., baseline, first, second, and third term) revealed no significant differences (F=1.94, df=15/534, α=.05). Table 5 shows no subject specialization subgroup mean teaching efficacy score differed significantly from that of any other at the baseline or at the end of any term.

For only the first subgroup, largely science majors, mean teaching efficacy scores declined significantly from the baseline to the end of the second term practicum, a finding that is difficult to explain in that it is unlikely that the practicum or earlier experiences would be in any way qualitatively different for this subgroup than for others.

A similar overall test of gender subgroup mean teaching efficacy scores and times indicated significant differences (F=4.05, df=7/543, α=.05). An examination of the confidence interval data showed no significant differences between the mean teaching efficacy scores of gender subroups at the baseline or the end of any term of the program. The mean teaching efficacy scores of female student teachers did not change significantly from term to term (F=2.54, df=3/226, α=.05). However, trend analysis revealed a significant *declining* linear trend (F=4.63, df=3/316, α=.05) in the mean teaching efficacy scores of male student teachers. Were these elementary male student teachers,

Table 5
Subject Specialization and Gender Mean Teaching Efficacy Scores: Baseline (B) through Term Three

	Term			
Subgroup	*B*	*1*	*2*	*3*
Math, Sci., Bus. Ed. Home Ec. & Ind. Ed.	26.56	25.19	24.00*	23.30*
English, Social Studies, and other languages	25.55	24.72	25.00	23.80
Art, Music, Theatre	26.77	25.00	23.93	26.00
Physical Education	28.00	23.80	24.33	24.23
Gender				
Female	26.92	25.80	24.14	25.50
Male	25.99	24.10	23.65	23.56

*Significantly different from the BASELINE score at the α=.05 level.
+Significantly different at the α=.05 level from a corresponding subgroup mean score recorded at the same time.

the finding might not have been surprising, since elementary teaching is still perceived as a more feminine than masculine occupation (Apple, 1987; Looker & McNutt, 1989; Rosenfeld & Stephan, 1978). Secondary teaching is, however, more gender neutral (Gottfredson, 1981), and is less likely to be viewed negatively (Brabeck & Weisgerber, 1989), hence finding a gender-based teaching efficacy difference is unexpected.

The personal teaching efficacy mean scores for subject specialization subgroups present a somewhat more positive picture. There were some significant increases (F=6.86, df=15/534, α=.05) as shown in Table 6.

Confidence interval data show, as indicated in Table 6, that at the baseline of the program the mean personal teaching efficacy score of subgroup four significantly exceeded that of subgroups one and two at the α=.05 level.

For the first two subgroups (75% of the students), mean personal teaching scores increased significantly during first term, although in not the second or third terms. Why personal teaching efficacy scores for these students did not increase during the extended practicum is puzzling. One would expect increases in personal teaching efficacy scores during an extended period of school experience, particularly if the experience is developmental and supervision is based on an adequate knowledge of effective teaching, focused on a manageable set of professional objectives and handled with sensitivity and genuine concern for the student (Housego & Boldt, 1985; Housego & Grimmet, 1983; Zimpher, 1987). The last two subgroups are small in contrast to the previously discussed ones and, although the changes in mean personal teaching efficacy scores are of similar magnitude, they fail to achieve significance.

The mean personal teaching efficacy scores for secondary gender subgroups are somewhat similar to the total group PREP Scale findings. There were significant (F=11.47, df=7/542, α=.05) personal teaching efficacy increases for both gender subgroups in first term, but for only female student teachers in the second term practicum. Other researchers report gender differences in the efficacy scores of practicing teachers favoring females (Anderson et al., 1987; Greenwood et al.,1990), and Ellis, Mathews, and Cieslak (1986),

Table 6
Subject Specialization and Gender Subgroup Mean Personal Teaching Efficacy Scores: Baseline (B) through Term Three

	Term			
Subgroup	*B*	*1*	*2*	*3*
Math, Sci., Bus. Ed. Home Ec., Ind. Ed.	35.52	39.67*	39.13	39.52
English, Social Studies, and other languages	34.71	38.13*	41.69	41.21
Art, Music, Theatre	36.23	39.53	42.00	42.10
Physical Education	40.62+	40.67	43.42	39.62
Gender				
Female	35.07	38.90*	41.75*	40.46
Male	36.14	39.81*	39.57	40.05

*Significantly different from the previous term score for the subgroup at the α=.05 level.
+Significantly different at the α=.05 level from a corresponding subgroup mean score recorded at the same time.

report beginning male teachers have been found to have greater difficulty adapting to the teacher's role.

Limitations

This study focuses only on student teachers' perceptions of preparedness to teach and efficacy feelings, not on actual teaching performances or pupil achievement. A student teacher's feeling prepared to teach and his or her self estimates of ability to effect changes through teaching or belief that such changes are attainable does not ensure that he or she *is*, therefore, well prepared to teach. A number of beginning teachers will be interviewed and questioned regarding their actual preparedness to teach in a subsequent phase of the work.

In this study, few student teachers felt unprepared to teach at any point in the program. Few responses were below three on the seven point Likert-style PREP Scale. Lower levels of the PREP Scale seem to have limited application.

In addition, because of the lack of random sampling and the lack of control, there are limitations concerning the generalizability of the findings of this study to other subjects in other locations under similar conditions at other times.

Conclusions

The research questions were answered as follows:

1. Secondary student teachers' feelings of preparedness to teach increased significantly in the first and the second but not the third term of the program.
2. All subject specialization subgroups increased significantly in feelings of preparedness to teach in first term; two of four subgroups did so in second term, and no subgroup did so in third term. Both gender subgroups increased significantly in feeling prepared to teach in the first and the second but not in third term.
3. No subject specialization or gender subgroup felt significantly more prepared to teach than any other at the baseline or at the end of any term of the program.
4. For secondary student teachers in total, there was a declining trend in teaching efficacy scores. They decreased significantly in first term. Personal teaching efficacy scores, by contrast, increased significantly in the first term but not beyond it.
5. The teaching efficacy scores of the first subject specialization subgroup declined significantly from the baseline score by the end of second term. This was not so for any other subject specialization or gender subgroup although the mean scores of males assumed a declining trend.

 The personal teaching efficacy mean scores of the first two subgroups increased significantly during first term only. No other subject specialization subgroup mean score increased significantly from one term to the next. Both male and female subgroups registered a significant increase in personal teaching efficacy in term one; only females continued growth in second term, and neither gender group showed growth in third term.
6. No subject specialization subgroup or gender subgroup mean teaching efficacy score differed significantly from that of any other at the outset of

the program or at the end of any term. At the baseline of the program, the mean personal teaching efficacy score of subject specialization subgroup four significantly exceeded that of subgroups one and two. There were no similar findings at the end of any term and no such findings regarding gender subgroup mean personal teaching efficacy scores.

Implications for Teacher Education

More than anything else, the findings of this study raise questions about the program's final term in which there were no significant increases in feelings of preparedness to teach or personal teaching efficacy for the total group or any subgroup, and teaching efficacy remained at a level significantly below that recorded when the program commenced. These findings are mirrored by the results of the author's investigation of the elementary teacher education program in which the content of the final term is similar (Housego, in press). Assuming the problem might be at least partially overcome through alterations to third term, what might be considered?

It has been suggested that the third term fails to capitalize on students' motivation to address the teaching and management problems they have experienced in the extended practicum. At that point, they are most able to identify their individual problems and, instead of addressing them, they move into the previously described largely theoretical term in which even the electives, having been chosen months earlier, may not meet current needs. As adult learners capable of framing their own objectives and pursuing what they value, student teachers are frustrated and unhappy about the expenditure of very limited time on course material in any way impractical or worse yet, repetitive. The educational psychology course could become a direct attempt to develop practical strategies for motivating students from varying backgrounds, to build a larger repertoire of management skills, and to develop and refine evaluation skills. Other courses could be scrutinized to see whether their content really meets the needs and expectations of soon-to-be-practicing teachers and whether there is enough commonality among course sections to assure students of comparable educational experiences and workloads. Perhaps some of the current third term content could be assigned as required to complete the Bachelor of Education degree and replaced with student-selected electives. It may be important to explore the relationships among the three terms and to assess the program as a whole, considering that problems that are manifested in third term may originate earlier in the program.

Also, it may be useful to reflect on the stage of development of student teachers after the practicum as a means of assessing their needs. Fuller and Bown (1975) proposed a sequence of student teachers' concerns on the basis of which four developmental stages were defined. In the first, a preteaching stage, students identify realistically with pupils but not teachers; in the second, their focus shifts to their own survival and the need to control the class and master content. The third-stage student teacher is concerned with actual teaching strategies while the fourth-stage student teacher is finally able to reflect on the needs of pupils and his or her own ability to relate to them as individuals. Whether the stages are distinct or overlapping (Burden, 1986, p. 192) is not clear. Assuming students who have completed the extended practicum to be probably at stage three and possibly at stage four and envisioning an appro-

priate final term of study, one would expect a preference for studying teaching strategies and learner characteristics rather than the so-called foundations of education.

Another set of implications grows out of the findings regarding the practicum, a term during which neither total group or subgroup teaching efficacy or personal teaching efficacy increased significantly, and during which female although not male student teachers' personal teaching efficacy scores increased. (Total group and gender subgroup PREP Scale scores did increase during the practicum). Some researchers (Katz, 1974) criticize the field experience component of teacher education programs as encouraging "excessive realism" among students, the tendency to model the classroom teacher's behavior instead of analyzing it as a basis for self-evaluation.

The quality of the educational experience of the practicum is uneven. Cooperating teachers are not all equally knowledgeable regarding effective teaching or to the same degree skilled in supervision. The same is true of university advisors. Nor are classrooms equally favorable as settings for the development of teaching strategies on the part of student teachers. Such problems are well known (Zeichner, 1990) and not easily solved; however, researchers continue to address these issues and make suggestions as to how field experience can be improved. Ideally, selection standards would be set for cooperating teachers and university advisors as opposed to placing students with volunteer teachers and providing few enticements for faculty to undertake supervision. Also, it would be helpful if there were opportunity to match student teachers and cooperating teachers and provide a workshop opportunity for some exploration of commonly held views. Certainly a mechanism for changing placements without prejudice is essential.

Further, although it is not so prevalent at the secondary as at the elementary level (Housego, in press), male student teachers seem to experience more difficulty than females. Women in this study increase significantly in personal teaching efficacy during the practicum. Males do not. Roberts and Nolen-Hoeksema (1989) report that, whereas women are influenced by positive and negative feedback and alter their self-evaluations accordingly, men take a more self-promotional stance and respond more to positive than negative feedback. Failure to process and act on negative feedback may help to explain male student teachers' limited increases in personal teaching efficacy. Certainly the practicum is an extended period of marginality (Clifton, 1979) and subservience, a time when it is difficult to be dominant, independent and oriented toward self-defined initiatives.

A final implication is derived from the declining teaching efficacy scores within the program. If one can attribute them to a confrontation with reality and a subsequent reduction of idealism, they can be understood. If they are viewed as a rapidly fading optimism concerning the credibility of teaching as a means of achieving desirable student learning goals, they are cause for concern, and methods of building a collegial spirit among students and between students and faculty in order to develop a more positive view of the profession should be sought.

References

Anderson, R., Greene, M., & Loewen, P. (1987). *Relationships among teachers' and students' thinking skills, sense of efficacy and student achievement*. Unpublished manuscript, University of Lethbridge.

Apple, M. (1987). Gendered teaching, gendered labor. In T.S. Popkewitz (Ed.), *Critical Studies in Teacher Education* (pp. 57-81.). New York: Falmer Press.

Armor, D., Conry-Oseguera, P., Cox, M., King N., McDonnell, L., Pascall, A., Pauly, E., & Zellman, G. (1976). *Analysis of the school preferred reading program in selected Los Angeles minority schools* (Report No. R-2007. LAUSD). Santa Monica, CA: Rand.

Ashton, P. (1984). Teacher efficacy: A motivational paradigm for effective teacher education. *Journal of Teacher Education, 35*(5), 28-31.

Ashton, P., Olejnik, S., Crocker, L., & McAuliffe, M. (1982, March). *Measurement problems in the study of teachers' sense of efficacy*. Paper presented at the meeting of the American Educational Research Association, New York.

Ashton, P., & Webb, R. (1986). *Making a difference*. New York: Longman.

Ashton, P., Webb, R., & Doda, N. (1983). *A study of teachers' sense of efficacy* (Final Report). Gainsville, FL: University of Florida.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.

Bandura, A. (1978). The self-system in reciprocal determinism. *American Psychologist, 33*, 344-358.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Berman, P., McLaughlin, M., Bass, G., Pauly, E., & Zellman, G. (1977). *Federal programs supporting education change. Vol. 7: Factors affecting implementation and continuation*. Santa Monica, CA: Rand.

Book, C., Byers, J., & Freeman, D. (1983). Student expectancies and teacher education: Traditions with which we can and cannot live. *Journal of Teacher Education, 34*(1), 9-13.

Brabeck, M., & Weisgerber, K. (1989). College students' perceptions of men and women choosing teaching and management: The effects of gender and sex role egalitarianism. *Sex Roles, 21*(11 & 12), 841-857.

Burden, P. (1986). Teacher development: Implications for teacher education. In J. Raths & L. Katz (Eds.), *Advances in Teacher Education* (vol. 2, pp. 185-219). Norwood, NJ: Ablex.

Cavers, L. (1988). *Teacher efficacy: Its relationship to school level organizational conditions and teacher demographic characteristics*. Unpublished doctoral dissertation, University of British Columbia.

Clifton, R. (1979). Practice teaching: Survival in a marginal situation. *Canadian Journal of Education 4*(3), 60-74.

Dembo, M., & Gibson, S. (1985). Teachers' sense of efficacy: An important factor in school improvement. *Elementary School Journal, 86*(2), 173-184.

Dewey, J. (1965). The relation of theory to practice in education. In M.L. Borrowman, (Ed.), *Teacher education in America: A documentary history* (pp. 140-171). New York: Teachers College Press, Columbia University. (Original work published 1904).

Ellis, J., Mathews, G., & Cieslak, M. (1986). *Differences in professional role performances and difficulties reported by first year male and female public school teachers*. DeKolb, IL: Northern Illinois University. (ERIC Document Reproduction Service No. ED 229-347)

Evans, E., & Tribble, M. (1986, April). *Perceived teaching problems, self-efficacy, and commitment to teaching among pre-service teachers*. Paper presented at the meeting of the American Educational Research Association, San Francisco.

Fuller, F., & Bown, O. (1975). On becoming a teacher. In K. Ryan (Ed.), *Teacher Education* (pp. 25-52). Chicago: National Society for the Study of Education.

Gibson, S., & Dembo, M. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76*(4), 565-582.

Glickman, C., & Tamashiro, R. (1982). A comparison of first-year, fifth-year and former teachers on efficacy, ego involvement, and problem solving. *Psychology in the Schools, 19*(4), 558-562.

Gottfredson, L. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counselling Psychology, 28*, 545-579.

Greenwood, G., Olejnik, S., & Parkay, F. (1990). Relationships between four efficacy belief patterns and selected teacher characteristics. *Journal of Research and Development in Education, 23*(2), 102-106.

Harootunian, B., & Yarger, G. (1981). *Teachers' conceptions of their own success: Current trends*. Washington, DC: ERIC Clearinghouse on Teacher Education.

Hoover-Dempsey, K., Bassler, O., & Brissie, J. (1987). Parent involvement: Contributions of teacher efficacy, school socioeconomic status, and other school characteristics. *American Education Research Journal, 24*(3), 417-435.

Housego, B.E.J. (1990a). A comparative study of student teachers' feelings of preparedness to teachers. *Alberta Journal of Educational Research, 36*, 223-239.

Housego, B.E.J. (1990b). Student teachers' feelings of preparedness to teach. *Canadian Journal of Education, 15*(1) 37-56.

Housego, B.E.J. (in press). Monitoring student teachers' feelings of preparedness to teach in a new elementary teacher education program. *Journal of Education for Teaching*.

Housego, B.E.J., & Boldt, W.B. (1985). Critical incidents in the supervision of student teaching. *Alberta Journal of Educational Research, 31*, 113-124.

Housego, B., & Grimmett, P. (1983). The performance-based/developmental debate about student teaching supervision: A typology and a tentative resolution. *Alberta Journal of Educational Research, 29*, 319-337.

Katz, L. (1974). Issues and problems in teacher education. In B. Spodek (Ed.), *Teacher education: Of the teacher by the teacher, for the child*. Washington, DC: National Association for the Education of Young Children.

Kelly, A. (1989). *An essay on the preparation of teachers*. Regina, SK: President's Office, University of Regina.

Lanier, J., & Little, J. (1986). Research in teacher education. In Wittrock, M. (Ed.), *Research on teaching* (3rd ed., pp. 527-569). New York: Macmillan.

Looker, E., & McNutt, K. (1989). Occupational expectations and educational attainments. *Canadian Journal of Education, 14*(3), 352-367.

Lortie, D. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.

Meijer, C., & Foster, S. (1988). The effect of teacher self-efficacy on referral chance. *Journal of Special Education, 22*(3), 378-385.

Midgely, C., Feldlaufer, H., & Eccles, J. (1988). *Change in teacher efficacy and student self and task-related beliefs in mathematics during the transition to junior high school*. Ann Arbor, MI: University of Michigan, Institute for Social Research.

Miller, P. (1988). *Teaching efficacy as an intervening variable in teacher's decisions about referring students for special education services*. Paper presented to the meeting of the American Educational Research Association, New Orleans.

Miklos, E., & Greene, M. (1987). Assessments by teachers of their preservice preparation programs. *Alberta Journal of Educational Research, 33*, 191-205.

Parkay, F., Olejnik, S., & Proller, N. (1988). A study of relationships among teacher efficacy, locus of control and stress. *Journal of Research and Development in Education, 21*(4), 13-22.

Poole, M., & Okeafor, K. (1989). *Teachers' interactions, personal efficacy and change implementation*. Paper presented to the meeting of the American Educational Research Association, San Francisco.

Roberts, T., & Nolen-Hoeksema, S. (1989). Sex differences in reactions to evaluative feedback, *Sex Roles, 21*(11 & 12), 725-747.

Rosenfeld, D., & Stephan, W. (1978). Sex differences in attributions for sex-typed tasks. *Journal of Personality, 46*, 244-259.

Tracz, S., & Gibson, S. (1986). *Effects of efficacy on academic achievement*. Paper presented to the meeting of the California Educational Research Association. (ERIC Document Reproduction Service No. ED 281 853)

Trentham, L., Silvern, S., & Brogdon, R. (1985). Teacher efficacy and teacher competency ratings. *Psychology in the Schools, 22*(3), 43-42.

Weinstein, C. (1988). Preservice teachers' expectations about their first year of teaching. *Teaching and Teacher Education, 4*(1), 31-40.

Woolfolk, A., & Hoy, W. (1988). *Efficacy beliefs and control orientations of prospective teachers*. Paper presented to the meeting of the American Educational Research Association, New Orleans.

Woolfolk, A., Rosoff, B., & Hoy, W. (1989). *Teachers' sense of efficacy and their beliefs about managing students*. Paper presented to the meeting of the American Educational Research Association, San Francisco.

Zeichner. K. (1990). Changing directions in the practicum: Looking ahead to the 1990s. *Journal of Education for Teaching, 16*(2), 105-131.

Zeichner K., & Tabachnick, B. (1981). Are the effects of university teacher education washed out by school experience? *Journal of Teacher Education, 32*(3), 7-11.

Zimpher, N. (1987). Current trends in research on university supervision of student teaching. In M. Haberman & J. Backus (Eds.), *Advances in Teacher Education* (vol. 3, pp. 118-150). Norwood, NJ: Ablex.

Alberta Journal of Educational Research *Vol. XXXVIII, No. 1, March 1992, 65-78*

Joe Fris
University of Alberta

Principals' Encounters with Conflict: Tactics They and Others Use

Fifteen principals of elementary and secondary schools in a moderately large school system in Western Canada provided detailed descriptions of nonroutine conflicts they had encountered. Protagonists used a variety of tactics in dealing with the conflicts. The principals most often attempted to manage a conflict by improving the shared information base, by censuring dysfunctional behaviors, and by managing organizational climate. Others' actions, as reported by the principals, included criticism, sabotage, and threats. The findings provide support for parts of Thomas' fivefold typology of conflict management behaviors. They also suggest that school principals prefer an integrative approach to conflict resolution, rather than a distributive approach.

Quinze directeurs d'écoles élémentaires et secondaires d'un système scolaire de taille moyenne au Canada Occidental ont fourni des descriptions détaillées de conflits importants auxquels ils avaient fait face. Les protagonistes utilisèrent des tactiques variées pour résoudre ces conflits. Les directeurs, le plus souvent, tentèrent de les résoudre en améliorant l'information commune, en censurant les comportements dysfonctionnels, et en gérant le climat de l'organisation. Les actions des autres, selon les directeurs, étaient la critique, le sabotage, et les menaces. Les résultats supportent certains aspects de la typologie à cinq éléments de Thomas des comportements de gestion du conflit. Ils suggèrent en outre que les directeurs d'écoles préfèrent une approche d'intégration à la résolution du conflit, à une approche distributive.

The Importance of Conflict

Conflicts have long been seen as important challenges in human affairs, primarily because of their negative effects. For example, Brickman (1974) defines conflict as "behavior by one party that injures or damages the interests of another party," and Deutsch (1973, p. 10), one of the most often quoted authorities, similarly characterizes conflict as incompatible activities that interfere with the accomplishment of objectives, cause injury, and reduce effectiveness. In the literature on educational organizations this negative perspective is also evident, but it is tempered consistently by acknowledgments of the potential for salutary effects. Further, conflict management is often linked to organizational effectiveness (e.g., Green 1988; Leithwood, Cousins, & Smith, 1990; Owens, 1987; Sergiovanni, 1987; Wynn & Guditus, 1984).

However, conceptualizations about conflict and its management (e.g., Owens, 1987) often come from studies conducted in such noneducational settings as private sector collective bargaining and national and international politics, and it seems that little of the information has been verified as ap-

Joe Fris is an associate professor in the Department of Educational Administration. His research focuses on conflict in schools.

plicable in school systems. Moreover, some of the findings from school studies are inconsistent with findings from noneducational settings. For example, Loewen (1983) found that the conflict management tactics used by school superintendents (chief executive officers) could not be categorized adequately when he relied on the typologies found in the literature: one of the nine types of conflict management tactics in the literature (appearing to lose) was not evident in his results; and four other types were not acknowledged in the literature: buying time, coalition destruction, environment control, and displaying naivety. Another inconsistency is evident in Boyd's (1989) study of how teachers deal with conflicts. The eight stages she identified in the development of conflicts do not match those found in the literature. Further inconsistencies emerge when Loewen's and Boyd's findings are compared with those of Abdennur. Whereas Abdennur (1987) found that social service volunteers tended to avoid conflict rather than confront or seek reconciliation, Loewen (1983) found that his school superintendents were quite aggressive when dealing with conflict, while Boyd (1989) found that teachers tend to seek reconciliation. Apparently typologies of coping behaviors may be occupation-specific.

It seems, then, that research is needed to validate (and extend) information about the management of conflicts in educational organizations. To this end, an investigation is underway to develop a substantial set of data (a projected 60 cases) concerning school principals' experiences with conflicts, and thereby to test the robustness of existing conceptualizations of conflict management. More specifically, the investigation is designed to (a) provide rich descriptions of important incidents in principals' experiences with conflict, (b) identify patterns in those experiences, and (c) relate the findings to existing theory.

At this time only the first 15 critical incident interviews with school principals that were conducted for Phase 1[1] of the four phase project have been subjected to an initial full analysis, but even these findings lend support to the proposition that the applicability in educational settings of extant information about conflict management needs to be tested. For example, one of the six questions that guide the investigation is "Who is involved and what causes conflict?" The related findings lend support to many of the propositions about conflict found in the literature (e.g., about complexity, likely protagonists, etc.) but also present some interesting anomalies (Fris, 1990); for example, that mainstream typologies (e.g., Corwin, 1969; French, Kast, & Rosenzweig, 1985; Sexton & Bowerman, 1979) have significant internal inconsistencies and may be unnecessarily complex. Indeed, at this stage it appears that a simple threefold typology of causes may account for most of the conflicts a school principal is likely to encounter:

1. Differences in values and goals: disagreements over what to do, the ends.
2. Incongruent role expectations: disagreements over how to proceed, the means. This class includes disagreements over how to organize a group of people for a collective enterprise, disagreements over the particular methods that various individuals will use to perform specific tasks, and disagreements over moral and ethical standards of behavior.
3. Disagreements over the distribution of limited resources: the rationing of the wherewithal. Included here are "shortages" in money, material and

equipment, time, space, status, influence (power and authority), and knowledge.

So far the data concerning types of conflicts indicate that the conflicts encountered by principals tend to center more on disagreements over how best to achieve certain objectives (type 2 above) and less on the distribution of scarce resources (type 3). Such findings, then, lend support to the notion that a critical analysis of information about conflict management is useful.

The purpose of this paper is to report findings that respond to the question "How do principals and others deal with conflicts?" and to relate those findings to pertinent research and theory. The findings are presented first from a *freeze-frame* perspective, then from a *time-lapse* perspective. Subsequently they are related to the models of conflict resolution behaviors advanced by Thomas (1976), Walton and McKersie (1965), and Axelrod (1984).

A Study of Principals

Participants

The first phase of the study was conducted in a Western Canada urban school system which comprised eight secondary schools (about 7,500 students) and 25 elementary schools (about 10,000 students). After the approval of senior administrators had been obtained, a description of the study was mailed to all principals in the system, and a purposive sample (principal's gender, elementary/secondary, size of school, socioeconomic status of community) was subsequently recruited by telephoning selected principals. The four women and 11 men who participated were the principals (and one vice-principal) of five larger (400-599 students) and four smaller (200-399 students) elementary schools, three larger (over 1,000 students) and three smaller (400-599 students) secondary schools. Their schools were located in both low and high socioeconomic communities. They provided interviews that lasted 60-90 minutes each.

Interviews

Each of the participants was asked to describe a school-related conflict that was "somewhat challenging" and not a routine problem. The interviews were recorded, with anonymity assured.

Data Analysis

The audiotapes of the interviews were transcribed to computer disks, a sample of the transcriptions was reviewed to check for accuracy, and each record was edited for readability. The length of the transcripts ranged from 11 to 25 single-spaced pages.

Each transcript was subjected to content analysis with the aid of a word processing program: in one window of the computer screen a transcript would be displayed and searched for references to pertinent phenomena—in this case any kind of action, maneuver, tactic, adjustment, and so on that seemed to be a move to deal with the conflict being described; such sections of the transcript would then be copied from the first window into another window, where they were grouped into categories of similar kinds of actions. This process, to use a film analogy, first distills the *plot* of an entire conflict incident into *snapshots* or *freeze-frames* of discrete actions, then creates *albums* (categories) of similar actions or developments. Because one of the principals had described two distinct conflict incidents, 16 such analyses resulted.

The full set of transcripts was also examined, to continue the analogy, for *recurring sequences of actions*: for recurring themes or *time-lapse sequences*. Each participant then received a copy of his or her interview transcript and the corresponding analysis for verification and further comment if desired. The full set of 16 analyses was then subjected to a further content analysis for cross-case consistencies.

Dealing with Conflicts: Freeze-Frame Perspectives

The interview transcripts contained 220 references to actions that were taken to deal with the conflicts; 146 of them described what the principals had done, 74 described what others had done.

Principals' Actions

The actions the principals took appear to fall into 11 categories. Only the first three are described in detail here because these types of actions were mentioned much more frequently than the others (Table 1).

Managing the Common Information Base

The principals most often attempted to deal with conflicts by working on the information that they *and others* used. This most often involved attempts to influence others' understandings. For example, a principal who described a conflict with members of a parents' committee (as well as teachers) over appropriate roles for parent volunteers took this action when, fairly late in the conflict, a trustee was about to become involved in the conflict:

I called up the trustee [who had been invited by the parents to their committee meeting] and I said, "You're being invited; go through the background."

Table 1
Conflict Management Strategies Used by the Principals

Major category *Subcategory*	*Number of principals*
1. Manage the information base (51 references)	
a. Influence others' understandings of the situation	8
b. Gather information	5
c. Verify information	4
d. Contradict	4
e. Impede the flow of information	4
f. Critique information	3
g. Try to change others' beliefs, actions	3
2. Censure unacceptable behavior (25 references)	
a. Set or reaffirm basic parameters	9
b. Put people in their places	3
c. Restrain, constrain	2
3. Manage climate (21 references)	
a. Provide moral support for others	8
b. Control the emotional pitch	3
c. Communicate particular values, model behavior	2
d. Emphasize importance of others' personal growth	2
e. Provide opportunity for saving face	1

Many of the other activities in this major category seemed to be directed at expanding the relevant information base, improving its quality, and disseminating its contents as widely as necessary. The following excerpts (with brief indication of the nature of the conflict) illustrate these dimensions:

Principal intercedes after two teachers get into a personal conflict. You know, you must communicate with your fellow principals if you have a problem. And some of the things I bounce off certain teachers who are on staff, who I can trust.

Two children, and families, in a running battle. I asked the teachers what [the boy] was doing. I spent many, many hours talking to the mother and to the father. I talked to many, many people to verify what I had heard.

Principal trying to curb photocopying by teachers. We informed the staff that photocopying was costing a lot of money. I kept tabs on it and I put this information in the bulletin that comes from the office every week.

Principal uncertain over best special placement for child. I talked to the mother in advance about the possibility of the special class placement. And the counselor took her over to look at the class and to meet the teacher. And she was invited to attend the meeting [for deciding special placements] if she would like to.

One sub-category in this category stands out as distinct from the other dimensions of information management, "impede the flow of information". Two examples of this kind of action appear in the following.

Principal and assistant superintendent disagree over appropriate role for the latter. I just quit telling the assistant superintendent about any problems. I told him nothing about the school. I handled this relationship by giving him no information. If he asked if anything was wrong in the school, I would talk about two or three things that were right. I encouraged him to use our meeting as an opportunity to share with me what he thought—and thus distracted him from his spying.

You can stop a conversation by standing up, using body language that makes the other person nervous, all of the tightening sorts of things that reduce the number of minutes of conversation.

A possible explanation is that this kind of action, even though it involves control of information, may in fact turn out to be part of the next category, censuring unacceptable behavior, with information control being a means rather than an end.

One phenomenon that has emerged with a fair degree of consistency (except in the above subcategory) is an apparent propensity to work problems out collaboratively, to involve the protagonists actively in the process of accommodating each competing interest as much, and *as equitably,* as possible. This common dynamic is illustrated in Table 2, a string of relevant excerpts from a case that involved a complex of discrete conflicts.

Censuring Unacceptable Behavior

The second most frequent category of responses to conflict had the general objective of reestablishing orderly, functional interpersonal relationships. For example:

Disagreement over appropriate role for assistant superintendent. I said, "[Only when] you begin to act in a confidential way can we discuss difficulties."

Table 2
Example of Collaborative Conflict Management

We had a new teacher, Mrs. Quick, a very fine master teacher. She had been a primary teacher and I put her in an intermediate classroom, a new experience for her.

Two parents, Mrs. Murphy and Mrs. Adams, became Mrs. Quick's parent volunteers. [Mrs. Murphy was President of the Parents' Committee and she and Mrs. Adams both had children in Mrs. Quick's class.]

Now, it got to the point where I could see that Mrs. Murphy wasn't helping out in the classroom: she was *teaching*. And two other people, the grade 3 teacher and the grade 4 teacher, had heard Mrs. Murphy teaching.

They didn't say anything to me but they talked with other staff members, who began to natter among themselves.

So I talked to Mrs. Quick and she said to me, "But she knows more [about computers] than I do; she can do what I'm getting her to do." So I said, "She is under your direction, yes?" And so I said, "We'll go with it." But the others weren't privy to that understanding.

Then, after I had been at a workshop for a couple of days, out of the school, when I came back in the school, Mrs. DeMuhn (who's in charge when I'm not here) said, "We've got a problem. While you were gone [teacher] Jones has been getting the staff together and really working them up. They don't like Murphy in the computer lab."

So then I went to Jones, and I got her to explain. And she was some upset: "She's ***! She's just ***! She's in here spying on us! She's walking in the halls. She's teaching that class; she's not in there as a parent!"

So we got everybody together as a staff because Jones, at that point, had them worked to a pitch.

And, you know, at this point teacher Quick was upset; she felt that nothing had been said directly to her but that they were all against her. I had given her support, she realized that but ... And both she and I realized there was only a couple more weeks left of the time the parent volunteers were going to come in and help us; we knew that we could just take that time and then it would be over.

But I had, you know, the situation with the staff. So we pulled them all together. I had two staff members lead it, the librarian and the person who is our teachers' association rep, brainstorm what their concerns were. And so they brainstormed ... they listed all their concerns.

People were on edge because this is not something people are used to doing. They weren't angry. They were, like, kind of ... They were kind of sitting back, like "I don't know whether I really want to be involved with this because this is going to get touchy and there's ... It may get hot, and I don't like this."

We went through all sorts of concerns that weren't related [to this issue] but, you know, they were just ... it was a way of breaking the ice. And finally they got down to the one I knew was crucial: that the principal wasn't being supportive.

I felt badly for teacher Quick because of course she's believing that she's caused all [the fuss]. She and I knew it was going to be over soon, and that I had been supportive of her. But she couldn't be seen to be a principal's pet, and she wanted to be accepted by the whole staff, because she was new to the school.

The solution they wanted was that Murphy be told she's got to be out of the computer class. And I said, "She's finished in only got a couple of sessions; we're going to ride that through."

Then they also wanted, because both Murphy and Adams were seen by the grade 3 teacher (and I think one other) as being in the halls too much, spying on them, that it be put in the newsletter that parents should report to the office before they come into the school, and that a sign be put up by the door: "Visitors please report to the office."

Those two suggestions I followed, even though I didn't agree with them. But I knew that you don't ask for suggestions and not follow them. So we let Murphy finish the last two helping sessions and that was the end of that [part of the troubles].

Grade 6 student a frequent truant, now also in trouble with the law. I said, "You're in grade 6, and I'm the boss."

[This student] would run—and he'd go! So what I used to do, if it looked like we were going to lose him, is take all his junk—his bag, his jacket, and all his personal belongings—and I'd take them and stash them somewhere where he couldn't find them. So when he ran, suddenly he didn't have all his worldly possessions with him [and he'd have to come back].

This kind of action, then, served to terminate behavior that prevented the principal from addressing the core issue.

Managing Climate

The third most common type of response seemed to focus on the improvement of others' emotional states. In the main, but not exclusively, this involved providing moral support:

Parent has complained to principal, apparently upset with teacher's treatment of daughter. I was very supportive of the teacher. I said, "If there is a problem, we will deal with the problem. I support you."

This kind of action, together with those in the other sub-categories seems to be directed at maximizing rationality and "empowering" others:

Controlling the emotional pitch. In preparation for the meeting with the parents we even generated and ran copies of a set of decision-making statements that we could pull out if we needed to—to help us remain very objective.

I acted quickly. I didn't want it to fester. I let them get through the affective stuff first. I decided that I was going to be nonjudgmental.

I tried mainly to be very calm and cool myself, to keep a fairly even tone and not be riled by him, and to take the line, "I am sorry you don't agree."

I am always very careful to think about how the other person can take my particular reply. I also try to keep it in a very light vein without getting too serious.

Communicating particular values. The reason I was doing it was to try to show the staff by my example.

I was able to develop a few economies here in the office; not much but a little bit to show them that the office was doing its bit.

Emphasize importance of others' personal growth. My main strategy is to create an environment in which people feel that they want to grow, they need to grow, they are excited about growth, they see it as a challenge, they see it as something they have a responsibility for. And to ensure that they can't use the excuse, "I want to grow but there is never any money or time".

Provide others with opportunities to save face. I suggested that we get together at the first opportune moment that day. And we were able to get some time that day. As we went through this whole discussion, I think it became apparent to me—and I think to them—that how the grade 7 teacher interpreted what he overheard being said and how the vice principal meant it were two different things. For example, when the vice principal said to the grade 7 teacher, "This other teacher called you a scumbag, and apparently she doesn't think you are treating her very well," he had put his own interpretation on [a conversation he had overheard]. And so the grade 7 teacher explained, "No. She's a friend of mine and she was teasing. We

insult each other regularly." So we went through some more of those kinds of things.

And then I was able to get them to focus a little bit on some of the strengths of each person—they did talk to each other about the things that they thought the other person did well. And it felt good.

Other Types

The other categories of actions for dealing with conflict that have emerged but remain to be confirmed as distinct types include using allies, protecting personal interests ("cover your butt"), focusing attention on the real problem, shifting responsibility for managing the conflict to whom it belongs, getting rid of the problem people, providing resources (time, specialist help), making concessions, and buying time.

Others' Actions

The principals reported a wide range of actions used by others, actions that represent perhaps as many as 16 categories (Table 3). Clearly using allies, criticizing, and crystalizing the problem were used most often.

A sense of what a school principal may reasonably expect to experience at the hands of others is communicated by the following excerpts.

Table 3
Types of Actions Taken By Others

Major Category *Subcategory*	*Number of principals*
1. Use allies (14 references)	
a. Individual	5
b. Group	4
c. Arbitrator	2
2. Criticize (11 references)	8
3. Crystalize the problem (8 references)	
a. Bring the issue into the open	2
b. State view	3
c. Question, test information	2
4. Express anger (7 references)	4
5. Gather information, spy (6 references)	4
6. Pre-empt or sabotage others' actions (5 references)	4
7. Threaten, intimidate (5 references)	4
8. Avoid/deny the problem (5 references)	3
9. Give up the fight (3 references)	2
10. Provide support (2 references)	2
11. Pester (2 references)	1
12. Protest, refuse (2 references)	1
13. Take control, remove control (1 reference)	1
14. Pin the blame on others (1 reference)	1
15. Refuse to be drawn into a conflict (1 reference)	1
16. Use misinformation, dissimulate, lie (1 reference)	1

Pre-empting the actions of others. Grade 6 student frequently a truant; now in trouble with law for breaking and entering, theft, etc.; but his diminutive stature garners lawyer's and judge's sympathy.
[The boy's assigned lawyer] said, "You're very angry and you shouldn't be in this business at all, dealing with children."
[The boy's assigned lawyer] didn't invite any of us to be witnesses, because "we were trying to get this kid put away." And she wouldn't talk to us; she was just being real snotty.
Threatening, intimidating: Union grievance over promotion of secretary. At the meeting before representatives of the board, there were two school trustees, a superintendent, ... a personnel manager, and me. And [on the other side] there were ... This was really ... I found this intriguing actually—there were the griever, the support person, the union rep, the president of the union, the negotiator for the [region], a support person, and another person who was learning the whole process. I mean, they just filled the other side of the table!
Avoiding or denying the problem: School has poor reputation, teachers resist principal's efforts to "turn the school (and them) around". When I stressed the importance of using teaching techniques that we know are effective, a typical reply from a teacher would be, "Yeah, but unless I feel good about it or unless the students feel good about me, unless there's lots of goodness, there will be no learning."
Refusing to be drawn into the conflict: Parents attempt to enlist help of trustee, to discredit the principal. The trustee listened and he said toward the end of it, "I have listened and I see that there have been meetings run that have not had consistent rules followed, that minutes have not been accurately kept. And it is completely understandable that a misunderstanding about the money took place. And I'm going to leave this meeting because I feel uncomfortable. You've made accusations against two of my employees, who are sitting here, and you now have to resolve that. I'm going to leave *you* to do that on your own."
Misinforming, lying: Parents attempt to discredit the principal. Then another letter went to the superintendent saying that the executive [of the parents' committee] had met and had decided ... Well, the executive hadn't met; only two of its members had met.

Dealing with Conflict: Time-Lapse Perspective

When the transcripts were examined for recurring sequences of discrete actions, one phenomenon occurred with sufficient frequency (12 of the 16 cases) to be noted here. This dynamic of conflict management involves a *shifting of gears*: typically a shift from a *softer* cooperative style to a *harder* forceful style. Table 4 provides an illustration of this dynamic.

Another aspect of this phenomenon is that in a few cases this dynamic included a reversion to the softer approach. Table 5 provides an example.

Discussion

The research reported here provides a basis for a critical examination of the information about conflict management that has been developed in arenas

Table 4
Example of a Shift from a Cooperative Style to a Forceful Style

When she came to me with her idea [for a show] I said, "Let's check it out with the staff."
She checked her idea out with the parents' committee.
[When one teacher came to me to object to the show] I just said, "Well I don't agree with you: and neither does the whole community." Because I had also done some research.
I called a staff meeting and said, "Look, we do not want recently healed tensions between staff members re-awakened, and we don't want to put our kids through that again." So we called off the show.
After we gave in to pressure and called off the show, a parent group came back to me and said, "We're not happy about this [decision to cancel the show]. We're going to take a petition and find out what the rest of the community, students, and the school committee thinks about it." About 99% wanted the show.
After the parents' survey showed great support for going ahead with the show, I called a staff meeting and I said, "Well, these people want us to do it; and they're paying our wages; and they have some say in this. So we're going ahead."
I told the objectors, "You don't have to help. And we will make a concession: we will not ***; we'll *** but we won't do what seems to be bothering you."
We checked with a group in *** that was reported to have these *** shows all the time. We found that they did not have the problem we were experiencing.
I decided I couldn't work under those conditions. So I covered myself by going to the superintendent and making sure he knew what was going on. And I said, "Look, you've got to transfer these people."
I covered myself by getting the backing of my parents.
I told the teachers concerned, "It would probably be best if you transferred."
The transfers came around May, June.

other than education. Hence some of the main findings are related now to other research and theory.

Thomas' (1976) widely used model of conflict posits that the behavior of protagonists is a function of two competing forces: a need to satisfy personal goals (assertiveness) and a need to maintain good relationships with other persons (cooperativeness). The model identifies five archetypal conflict behaviors (Cooze, 1990, p. 23):

1. Forcing: high on assertiveness, low on accommodativeness.
2. Accommodating: high on accommodativeness, low on assertiveness.
3. Collaborating: high on both imperatives.
4. Avoiding: low on both imperatives.
5. Compromising: moderate on both imperatives.

In light of this model, the 16 cases analyzed in this investigation suggest three important things about principals' management of conflicts. First, the principals tended to treat the conflicts as problems to be solved collaboratively, the third approach in the Thomas matrix. This suggests a preference for the *win-win* approach to conflict management rather than the highly aggressive *win-lose* approach. In Walton and McKersie's (1965) terms, this is a preference for *integrative bargaining* as opposed to *distributive bargaining*. Integrative bargaining is possible when the objectives of the parties can be integrated to some degree: "Integrative potential exists when the nature of a problem permits solutions which benefit both parties, or at least when the gains of one party do not represent equal sacrifices by the other" (p. 5). This is closely related to what

Table 5
Example of a Cooperative-Forceful-Cooperative Shift

The parents' committee were all excited about the computer lab and they wanted software. They suggested that money from fundraising be used to get the software. And I said to them, "Well, I've just put an order in, to be paid from school funds. But would you like to use the fundraising money to pay for that?" And they said, "Oh, yes, yes, yes." So I made the change right after the meeting, in front of most of them.

Shortly thereafter, Brenda, President of the Parents' Committee, came in and said to my secretary, "I want the books, the books on fundraising." At that point we were managing all of the fundraising money. And the fundraising accounts were always open to them. But in this book we also kept records of my principal's account. So I said, "No, the book can't leave the school," and that if she wanted to take a look at it she could take a photocopy of the fundraising page or look at it here. And I said, "Would you like to look at it here?" "No, no, no. I want to take the accounts home."

I got a signal right there: her husband is a banker. And so I said to her, "Is there a problem, Brenda?" And she said, "You've spent [parents'] money. You spent money on software." And I said, "Yes, but that's what the committee decided." And I also said, "Now, perhaps this was a mistake. If you're not comfortable with that I can reverse that."

Now that was a mistake. I should have said, "Well, we should discuss that at the next meeting." But seeing her in distress, and knowing that she had a bee in her bonnet and was being guided by the background person, I thought, "Oh my god, here I go again. We're into another." And to me it was neither here nor there; we had school money to pay for it.

Well, she then went to a trustee, who was a first-year trustee and so was getting actively involved and listening to her. And then I got a copy of a letter she'd sent the superintendent, again accusing me of mismanagement of funds.

Well, because that was a very serious accusation, I went to the assistant chairperson of the Parents' Committee, Roberta, and alerted her: "I think we have a problem." And I outlined what had gone on and she said, "There's no way she can say that." She had been at the meeting.

Then Brenda decided that she would call an executive meeting without me present. But Roberta, when she got the call, called me and said, "This is what's going on and I'm not going to allow the meeting to go on." So she refused to go and got some of the other women to boycott it too.

But then another letter went to the superintendent saying, "The executive has met and has decided ..." even though the executive hadn't met: only two ladies had met. And they wanted a formal audit done of the books and they wanted to set up a separate account.

At that point I had begun to realize that the woman was not ... was being manipulative, was mentally ill ... was emotionally distressed. And the assistant superintendent's response was, "Call up her husband and have her husband talk to her." Well, I couldn't see that as a solution, I didn't see that as an alternative.

So I talked with Roberta. And she said, "At the next executive meeting we're going to address this."

In the meantime I heard that Brenda had invited the trustee to come to our next executive meeting. So I called him up and I said, "You're being invited; go through the background." And when we met before the meeting, he said, "What do you expect from this? And how can I help?" He realized that the woman was under much emotional distress, and he realized the game that was being played.

At the meeting Roberta, the Vice-President, just outlined what had happened: du, du, du, du! And at that point Brenda was almost shattered. Interestingly, the person who had been prompting Brenda behind the scenes did not show up at that meeting; she left Brenda to deal with all of the mess herself. The trustee listened and he said toward the end of it, "You know, I have listened and I see that there have been meetings run that have not had consistent rules followed, that have not had accurately kept minutes. And it is completely understandable that misunderstanding about the money took place. And I'm going to leave this meeting because you have made accusations against two of my employees, and *you* now have got to resolve that. I'm going to leave you to do that on your own."

Table 5 (continued)

Then we talked more about it. In fact I gave support to the President, because they were ready to ask for her resignation at this point and I don't know how she kept herself together. She was literally shattered. And I said, "Let's leave it a month until all has calmed down." We revised the constitution, we decided to always follow Roberts' Rules of Order, and that the meetings would always be taped.

But although we had that worked out, I still had to get rid of Brenda as a parent volunteer [who was seen by the teachers as a spy and meddler]. So I asked for a meeting with her and said that I'd really appreciated the help that she had given to the school. And, you know, it started on a really positive note, with me specifically recalling some of the [good] work she had done. But I went on to tell her that I, and the staff, had some concerns about the number of parents that were in the school all the time and that I was going to have to ask her not to continue as a parent volunteer because there was some ... that staff were having some concerns about the number of parents that were in school and about her in particular. She began to cry, and cry, and cry, and took it as if she had done something terribly wrong.

After that she made an appointment to see me, and I asked the assistant superintendent to be there. So she and her husband came in to see me. She was under great emotional distress and he handled the interview. And the husband said, "I think we've given a great deal to this school. I think ... we think the school may be closing. We think it might be a better idea if we transferred to another school." So he was suggesting to me he'd like to get his wife out of this setting. And we talked about that and agreed that would be a good idea. He knew that she was under distress. And he was trying to engineer a change, to let her have dignity. She was relieved to have him take over. I was relieved that they were going, transferring both children to another school. And we left it at that.

game theorists call the varying-sum game. Distributive bargaining, in contrast, occurs when the attainment of one party's goals is in basic conflict with those of the other party and one party's gain is a loss to the other (pp. 4-5), what game theorists refer to as a fixed-sum situation.

Second, the strategies used suggest that *attitudinal restructuring*, adapted from Walton and McKersie (1965), is a favored conflict management tool. Attitudinal restructuring activities are those that have the object of influencing relationships between parties, including competitiveness and cooperativeness. This objective seems to underlie the two types of actions that were reported most often in this study: managing the information base and managing climate.

Finally, the interview records point to the importance of being prepared to change a management style as the circumstances of a conflict change. Some of the scenarios suggested that an advance in conflict management is sometimes achieved by abandoning a collaborative or integrative approach and switching to *forcing* tactics. Forcing tactics were evident in *contradicting* (Type 1d actions in Table 1) and *censuring of unacceptable behavior* (Type 2 actions in Table 1). They were also evident in some of the other tentative classes of conflict management behaviors: using allies, safeguarding personal interests, shifting responsibility for a problem to whom it belongs, and attempting to get rid of problem people. This general pattern is consistent with Content's (1986) conclusion that suburban California principals most often use compromising and collaboration, followed by avoiding, accommodating, and forcing. A more important point, however, is that success in conflict resolution has been linked consistently (Axelrod, 1984) with the ability to shift from cooperation, to forcing, and back, the *tit-for-tat* strategy:

1. Start off *nice*. Never be the first to aggress.
2. *Retaliate* when your opponent aggresses.
3. Always be *forgiving*. If your opponent is nice or cooperates, forget the past and return the nice move.
4. Be very *clear and open about your strategy*: let others know that these are your principles.

The cases on which this report is based seem to support the contingency principle (Tosi & Carroll, 1976) that underscores the tit-for-tat approach.

These three general findings concerning principals' ways of dealing with conflict provide support for parts of the Thomas (1976) model of conflict management behaviors. However, it should be kept in mind that the findings may have been influenced by such factors as the participants electing to describe *safe* or flattering incidents. This might explain, among other things, the contrast between these somewhat gentle findings and Loewen's (1983) depiction of school superintendents (CEOs of school districts) as primarily assertive, if not aggressive, in the management of conflicts. Another consideration is that the way principals manage conflict reflects their philosophies of administration: which raises the kinds of philosophical issues that Hodgkinson (1983) discusses. For example, the tactic *get rid of problem people* clearly relates to the issues associated with the "excisionistic fallacy". But such matters are beyond the scope of this article.

Note

1. In Phase 2, "cultural context" was varied: 18 school principals in New Zealand were interviewed. In Phases 3 and 4, the cultural context will be changed again.

References

Abdennur, A. (1987). *The conflict resolution syndrome: volunteerism, violence, and beyond*. Ottawa, ON: University of Ottawa Press.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic.

Boyd, T.J. (1989). *School staff conflict: The teacher's experience*. Unpublished research project report, University of Alberta.

Brickman, P. (1974) *Social conflict*. Lexington, MA: Heath.

Content, S.H. (1986). *Conflict management styles of principals in elementary and secondary schools*. Unpublished doctoral dissertation, University of California, Los Angeles.

Cooze, J. (1990). Conflict resolution strategies. *Canadian School Executive, 9*(8), 22-24.

Corwin, R.G. (1969). Patterns of organizational conflict. *Administrative Science Quarterly, 14*, 504-520.

Deutsch, M. (1973). *The resolution of conflict: constructive and destructive processes*. New Haven, CT: Yale University Press.

French, W.L., Kast, F.E., & Rosenzweig, J.E. (1985). *Understanding human behavior in organizations*. New York: Harper & Row.

Fris, J. (1990, November). Principals' encounters with conflict: Types, related factors, effects. *New Zealand Journal of Educational Administration, 5*, 41-49.

Green, W.J. (1988). *An analysis of the tasks, skills, and personal characteristics associated with the role of school superintendent*. Unpublished doctoral dissertation, University of Alberta.

Hodgkinson, C. (1983). *The philosophy of leadership*. Oxford, UK: Blackwell.

Leithwood, K.A., Cousins, J.B., & Smith, M.G.R. (1990). The principals' problem solving: Types of problems encountered. *The Canadian School Executive, 9*(9), 18-22.

Loewen, R.D. (1983). *Tactics employed by senior educational administrators when engaged in conflict*. Unpublished doctoral dissertation, University of Alberta.

Owens, R.G. (1987). *Organizational behavior in education*. Englewood Cliffs, NJ: Prentice-Hall.

Sergiovanni, T.J. (1987). *The principalship: A reflective practice perspective*. Newton, MA: Allyn and Bacon.

Sexton, M.J., & Bowerman, K.D. (1979) Conflict handling for secondary school principals. *NASSP Bulletin, 63* (April), 1-3.
Thomas, K.W. (1976). Conflict and conflict management. In M.D. Dunnette (Ed.), *Handbook of industrial and organizational psychology*. Chicago, IL: Rand McNally.
Tosi, H.L., & Carroll, S.J. (1976). *Management: contingencies, structures, and process*. Chicago, IL: St. Clair Press.
Walton, R.E., & McKersie, R.B. (1965). *A behavioral theory of labor negotiations*. New York: McGraw-Hill.
Wynn, R., & Guditus, C.W. (1984). *Team management: Leadership by consensus*. Columbus, OH: Merrill.

For Contributors

Please prepare manuscripts in accordance with these guidelines.

Format

- Manuscript in triplicate.
- Typewritten or printed, double-spaced throughout.
- Abstract of approximately 100 words on a separate sheet.
- Manuscript not to exceed 6,500 words excluding graphics.
- Author's name and affiliation on a separate title page; title only on first page of manuscript to ensure anonymity in the review process.

Referencing

- Sources in parentheses after each reference (direct or otherwise), giving author's name (unless mentioned in text), year of publication, and page number(s).
- Direct quotes of 40 words or less in double quotation marks incorporated into text; those longer than 40 words in indented block format.
- All sources listed alphabetically at the end of the manuscript under the heading *References* using the APA style.
- Explanatory notes, numbered consecutively and marked in the text with superscript numerals, may appear before the References under the heading *Notes;* any citations in notes follow the same format as other references.

Graphics

- Each table and figure on a separate page at the end of the manuscript.
- Figures, charts, and diagrams in camera-ready form.
- Number tables in Arabic numerals and center over each heading, e.g., *Table 1.*
- Number figures in Arabic numerals before the caption at foot of each graph, e.g., *Figure 1: ...*
- Indicate placement of figures and tables in text, i.e., *Insert Table 3 here.*

Style

- *AJER*'s editorial style conforms closely to the *Publication Manual of the American Psychological Association.*
- Editorial changes may be made to manuscripts.
- For spelling, consult *Webster's New Collegiate Dictionary;* spelling in quoted material remains as in the original.

Diskettes

If manuscript is accepted for publication and has been prepared using a Macintosh, IBM, or IBM compatible microcomputer, please send diskette. Microsoft Word, MultiMate, WordPerfect, WordStar, Xerox Writer, XyWrite, and DisplayWrite 3 **only** are acceptable.
Important: Please indicate program used.

Submission to *AJER* implies the manuscript is not being considered for publication elsewhere.

AJER is indexed in the *Alberta Education Index, Canadian Education Index, Current Contents/Social and Behavioral Sciences, ERIC/Current Index to Journals in Education,* and *Social Science Citation Index;* appropriate articles are abstracted in *Cognitive Development Abstracts, Educational Administration Abstracts, Educational Technology Abstracts, Family Abstracts, Language Behavior Abstracts, Multicultural Education Abstracts, Psychological Abstracts, Research into Higher Education Abstracts, School Organization and Management Abstracts,* and *Sociology of Education Abstracts,* and reviewed by the *Journal for Research in Mathematics Education.*

If undelivered return to:

The Alberta Journal of Educational Research
4-116 Education North
University of Alberta
Edmonton, AB
T6G 2G5

Publications Mail Registration Number 1436



THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XXXVIII NUMBER 2 JUNE 1992

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

The Alberta Journal of Educational Research

Published in March, June, September, and December
by the Faculty of Education, University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation, and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Editor: Robert H. Short

Consulting Editors

James Chapman
Massey University

Rod Clifton
University of Manitoba

Dianne Common
Lakehead University

Carol Crealock
University of Western Ontario

John Hattie
University of Western Australia

Billie Housego
University of British Columbia

Linda Larocque
University of Alberta

Stewart J.H. McCann
University College of Cape Breton

Les McLean
Ontario Institute for Studies in Education

Linda Phillips
Memorial University of Newfoundland

Sharon Robertson
University of Calgary

Faculty Advisory Committee

John G. Paterson (Chair)
Fern Snart
Joyce Edwards
Graham Fishburne

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada, from the Alberta Advisory Committee for Educational Studies, and from the Office of Research Services, University of Alberta.

The subscription rate is $32.00 per year for individuals, $40.00 per year for institutions. Add $8.00 for delivery outside Canada. Single copies are $8.00 each. Subscriptions and sales in Canada will be charged 7% GST. Please make cheques payable to *The Alberta Journal of Educational Research*. Back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.

Publications Mail Registration Number 1436

Faculty of Education
University of Alberta

The Alberta Journal of Educational Research

Volume XXXVIII, Number 2, June 1992

Christopher Wilson
Leonard L. Stewin **79** Semantic Differential Responses to Educational Posters on Acquired Immune Deficiency Syndrome (AIDS)

James Ryan **91** Formal Schooling and Deculturation: Nursing Practice and the Erosion of Native Communication Styles

Margaret Haughey
L. MacElwain **105** Principals as Instructional Supervisors

Douglas J. Smith
Robert P. Sanche **121** Saskatchewan Interns' Concerns at Three Stages of a Four-Month Practicum

Murray Sandell
Keith Sullivan **133** Teacher Disillusionment and Supervision as a Part of Professional Development

Reg Fleming **141** Teachers' Views of Technology

ISSN 0002-4805

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 2, June 1992, 79-89*

Christopher Wilson

and

Leonard L. Stewin
University of Alberta

Semantic Differential Responses to Educational Posters on Acquired Immune Deficiency Syndrome (AIDS)

The responses of 131 undergraduate university students, on a nine-item semantic differential scale, to eight educational posters dealing with the Acquired Immune Deficiency Syndrome (AIDS), were tested using Analysis of Variance (ANOVA). Post hoc analysis, using the Scheffé test, was performed in order to determine which differences in ratings were significant. The posters utilized different themes and messages ranging from highly informative to humorous. Some relied heavily on homosexual content. The common objective of the posters was to educate individuals about AIDS and AIDS prevention. Two posters were consistently rated as more informative, reassuring, effective, decent, and better than the others. The first utilized an objective and informative approach, whereas the second stressed personal responsibility in disease prevention. A number of posters were negatively received. Gender differences in reactions were not noted. Implications for educational programs and further research are discussed.

Les réponses de 131 étudiants du premier cycle universitaire aux neuf items d'une échelle de sémantique différentielle, comportant huit affiches éducatives sur le Syndrome Immuno-Déficitaire Acquis (SIDA) ont été évaluées à l'aide de l'analyse de la variance. L'analyse post-hoc, à l'aide du test de Scheffé, a été faite pour déterminer les différences significatives dans les réponses. Les affiches portaient sur des thèmes différents et des contenus allant des plus informatives aux plus humouristiques. Certaines affiches s'inspiraient d'un contenu d'homosexualité. Les affiches avaient pour objectif commun d'éduquer les individus au sujet du SIDA. Deux affiches étaient jugées, la plupart du temps, comme étant plus informatives, plus rassurantes, plus efficaces, plus décentes, et meilleures que les autres. La première utilisait une approche objective et informative, alors que la seconde mettait l'accent sur la responsabilité personnelle dans la prévention de la maladie. Certaines affiches ont eu une réception négative. Il n'y avait pas de différences dans les réactions dues au sexe des répondants.

The Acquired Immune Deficiency Syndrome (AIDS) has the potential to become the greatest natural tragedy in human history (Masters, Johnson, & Kolodny, 1988). Health educators have adopted a tripartite model to deal with the problem (Shernoff & Palacios-Jimenez, 1988), focusing on primary, secondary, and tertiary prevention. Effective education has been advocated as the

Christopher Wilson is a doctoral student in educational psychology. His research interests include measurement of personality, attitudes, and intelligence.
Leonard Stewin is a professor in the Department of Educational Psychology. His research interests focus on the developmental aspects of concept formation, information processing and its applications to learning, and the social psychology of the classroom.

only viable means to prevent, reduce, and delay the spread of AIDS (Lenaghan & Lenaghan, 1987; Norman, 1986) and occurs at the level of primary prevention. Posters, leaflets, pamphlets, radio, and television advertising have been part of numerous campaigns directed at communities all over the world in attempts to control the disease (Fineberg, 1988a; 1988b). However, although ignorance about AIDS may be eradicated, this is only half the battle. Evidence of a study of knowledge, behavior, and attitudes of 38,000 Canadian youths indicates levels of knowledge alone have not brought about changes in sexual practices (King et al., 1988). Baldwin and Baldwin (1988a), Fineberg (1988a), Flora and Thoreson (1989), Kegeles, Adler, and Irwin (1988), Kelly, St. Lawrence, Hood, and Brasfield (1988), Sherr (1987), and Strunin and Hingsen (1987) found similar gaps between knowledge about AIDS and behavior change.

Spurgeon (1988) states that the Australian National Advisory Committee on AIDS found a number of barriers to absorption and behavior change in AIDS education. These included messages that were too reassuring, paranoia resulting from too little reassurance, a negative image of condoms, and a perception that condom use was an acknowledgment of guilt. Baldwin and Baldwin (1988b) note that little research has been undertaken to evaluate what kinds of education will be appropriate. Flora and Thoreson (1989) maintain current research contributes little to the understanding of how to tailor messages and utilize appropriate methods of communication in AIDS education.

In the present study reactions to different educational posters used in public health campaigns were measured according to a nine-item semantic differential scale in order to provide feedback about how the subjects perceived the educational posters. Ratings on the semantic differential scale are intended to reveal particular dimensions that people use to qualify their experience and may reflect intensity of feeling for an attitude object (Zimbardo & Ebbeson, 1969). These particular dimensions may be of relevance to AIDS educational programs, particularly in light of Fisher's (1980; 1986) and Fisher, Byrne, White, and Kelley's (1988) findings that suggest emotional orientations toward sexuality influence an individual's ability to learn sexual material.

Method

Subjects

One hundred and thirty-one third- and fourth-year students in a Canadian university faculty of education program took part in the survey. There were 56 male and 75 female students in the sample. The mean age of respondents was 23.7 (SD=5.61) and ranged from 19 to 57 years. Twenty-six were married (8 male, 18 female) and 105 single (48 male, 57 female). The students were tested in a classroom setting during regular instructional time. Participation was entirely voluntary.

Survey

Contrasting informational and educational posters on AIDS that relied on different themes to convey information were surveyed. A brief description of each poster follows.

The first poster showed a young male wearing denim jeans, viewed from the back. There is a condom in his pocket. The poster bears the inscription "A

rubber is a friend in your pocket." The second poster depicts two muscular male chests touching each other, with the statement "Condoms—Use Them." The third poster is captioned "Which of these will give you AIDS?" and shows misconceptions concerning possible means of transmission, that is, through using cutlery, toilets, drinking glasses, or telephones. The fourth poster shows a male reclining provocatively, in a state of undress in bed, and is captioned "You'll never forget the feeling of safe sex." Below that inset into a little box is the statement "Hot Nights Safe Sex," advocating the use of condoms. The fifth poster shows a picture of Michaelangelo's statue *David* with the caption "Get It Under Cover." Inset into a box below is the warning to "Play it safe." The sixth poster shows a well proportioned muscular upper body of a male with a caption suggesting that it is masculine or manly to use a condom to prevent transmission of diseases. The seventh poster shows a group of males with their arms around each other at the beach, wearing brief swimsuits. The caption reads "Have a safe sex summer." The final poster shows a picture of a pensive adolescent male staring at the viewer. The caption reads "Stopping AIDS is up to you."

Respondents were asked to rate their responses to these eight posters according to a nine-level semantic differential scale. The semantic differential scale was developed by Osgood as a method of studying attitudes by focusing on the meaning that people give to a word, concept, or symbol (Maguire, 1973).

Results

Differences in reactions to all posters along each of the nine adjectives of the semantic differential scale displayed in Table 1 were tested using Analysis of Variance (ANOVA). Post hoc analyses, using the Scheffé test, were then performed in order to determine which differences in pairs of means were significant.

The first adjective of the semantic differential required the subjects to rate the posters along a continuum in an attempt to estimate how informative they were. There were significant differences between mean ratings of the posters ($F = 46.92$, $p<.001$). Post hoc comparisons revealed that two posters were rated more informative than the rest. The most informative poster, according to the respondents, was clearly the poster that showed ways in which individuals

Table 1
Semantic Differential Items Used in the Study

Informative	Uninformative
Emotive	Unemotive
Suggestive	Persuasive
Reassuring	Threatening
Effective	Ineffective
Offensive	Inoffensive
Good	Bad
Moral	Immoral
Decent	Indecent

Table 2
Significant Differences Between Mean Semantic Differential Ratings
Adjective—Informative/Uninformative

Mean	*Poster*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
3.56	1	- -							
3.52	2	NS	- -						
1.45	3	*	*	- -					
3.73	4	NS	NS	*	- -				
3.72	5	NS	NS	*	NS	- -			
3.60	6	NS	NS	*	NS	NS	- -		
3.84	7	NS	NS	*	NS	NS	NS	- -	
2.90	8	*	*	*	*	*	*	*	- -

*pairs of means significantly different at p< .05 level

could avoid contracting AIDS and moreover attempted to dispel misconceptions about the disease. The eighth poster showed a very timid looking adolescent staring at the viewer and was captioned "Stopping AIDS is up to you." It is possible that this poster was considered informative because it stressed personal responsibility in preventing the disease. No other pairs of posters were rated as being significantly different. The results appear in Table 2.

The second semantic differential item required the sample to rate posters according to whether they felt that they were emotive or not. There were significant differences between mean ratings of the posters (F=8.1, p<.001). Post hoc comparisons revealed that Poster 2, which showed two naked male chests touching each other, was considered to be more emotive than the posters emphasizing an informative approach, stressing personal responsibility or utilizing males, either individually or in groups, to promote condom use. The poster stressing personal responsibility in AIDS prevention was rated less emotive than the poster of a male reclining provocatively in bed urging people to use "safe sex." No clear differences emerge, but it appears the somewhat blatant sexual theme of Poster 2 may have prompted the relatively extreme evaluations. Explicitly depicted sexual contact is absent from the remaining posters. The majority of mean responses were not significantly different. The results appear in Table 3.

The third item of the semantic differential required individuals to rate the posters along a continuum as to how suggestive, as opposed to persuasive, they were. There were significant differences between mean ratings of the posters (F= 61.74, p<.001). Post hoc comparisons revealed that the third poster, which depicted how one could not contract the disease was rated as more persuasive than all the remaining posters. Poster 8, which showed an adolescent staring at the viewer and stressing personal responsibility in stopping AIDS, was rated as more persuasive than posters 2 and 4, which had homosexual themes, but was not considered as persuasive as the highly informative posters. Conversely, Poster 4, which depicted a semi-clad male in bed stressing the sensuality of safe sex, was rated as more suggestive than all the

Table 3
Significant Differences Between Mean Semantic Differential Ratings
Adjective—Emotive/Unemotive

Mean	*Poster*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
3.20	1	- -							
2.64	2	NS	- -						
3.60	3	NS	*	- -					
2.64	4	NS	*	NS	- -				
3.01	5	NS	NS	NS	NS	- -			
3.03	6	NS	NS	NS	NS	NS	- -		
3.32	7	NS	*	NS	*	NS	NS	- -	
2.91	8	NS	*	NS	NS	NS	NS	NS	- -

* pairs of means significantly different at p< .05 level

posters except for the one showing naked male chests. The results appear in Table 4.

The fourth item of the semantic differential required individuals to rate the posters along a continuum as to how reassuring, as opposed to threatening, they were. There were significant differences between mean ratings of the posters (F=79.66, p<.001). Post hoc comparisons revealed that Poster 3 was consistently rated as more reassuring than all the remaining posters. Poster 8 was rated as more reassuring than all the remaining posters except for that stressing the masculinity of using condoms (Poster 6). Poster 3 relied on a purely informative approach, whereas Poster 8 stressed personal responsibility. Poster 2, which depicted two unclothed males bodies, was rated more threatening than reassuring than all the posters except for Poster 7, which showed a group of males in bathing suits with their arms around each other and advocating use of the condom. The results appear in Table 5.

Table 4
Significant Differences Between Mean Semantic Differential Ratings
Adjective—Suggestive/Persuasive

Mean	*Poster*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
2.79	1	- -							
2.16	2	*	- -						
4.41	3	*	*	- -					
1.91	4	*	NS	*	- -				
2.40	5	NS	NS	*	NS	- -			
3.16	6	NS	*	*	*	*	- -		
2.64	7	NS	NS	*	*	NS	NS	- -	
4.11	8	NS	*	*	*	NS	NS	NS	- -

*pairs of means significantly different at p< .05 level

Table 5
Significant Differences Between Mean Semantic Differential Ratings
Adjective—Reassuring/Threatening

Mean	*Poster*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
3.41	1	- -							
4.40	2	*	- -						
1.82	3	*	*	- -					
3.60	4	NS	*	*	- -				
3.82	5	NS	*	*	NS	- -			
2.87	6	*	*	*	*	*	- -		
4.02	7	*	NS	*	NS	NS	*	- -	
2.75	8	*	*	*	*	*	NS	*	- -

* pairs of means significantly different at p< .05 level

The fifth item of the semantic differential required individuals to rate the posters along a continuum as to how effective the various posters were. There were significant differences between mean ratings of the posters (F=35.00, p<.001). Post hoc comparisons revealed that two of the posters were consistently rated more effective than the rest. The third poster in particular, which relied on a highly informative approach, was rated the most effective in comparison to all the other posters. The eighth poster, which emphasized personal responsibility, was also rated as more effective than all the remaining posters, except for Poster 6, which stressed the masculinity of using condoms. Poster 7, which showed a group of males in swimsuits was considered less effective than the poster that showed only one male and emphasized the masculinity of condom use. The results appear in Table 6.

The sixth item of the semantic differential required individuals to rate the posters along a continuum as to how offensive or inoffensive they were. There were significant differences between mean ratings of the posters (F=111.47,

Table 6
Significant Differences Between Mean Semantic Differential Ratings
Adjective—Effective/Ineffective

Mean	*Poster*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
3.69	1	- -							
3.84	2	NS	- -						
1.90	3	*	*	- -					
3.68	4	NS	NS	*	- -				
3.53	5	NS	NS	*	NS	- -			
3.22	6	NS	NS	*	NS	NS	- -		
3.87	7	NS	NS	*	NS	NS	*	- -	
2.64	8	*	*	*	*	*	NS	*	- -

*pairs of means significantly different at p< .05 level

Table 7
Significant Differences Mean Semantic Differential Ratings
Adjective—Offensive/Inoffensive

Mean	*Poster*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
4.08	1	- -							
2.40	2	*	- -						
5.42	3	*	*	- -					
3.09	4	*	*	*	- -				
3.17	5	*	*	*	NS	- -			
4.57	6	NS	*	*	*	*	- -		
3.18	7	*	*	*	NS	NS	*	- -	
5.43	8	*	*	NS	*	*	*	*	- -

*pairs of means significantly different at p< .05 level

$p<.001$). Post hoc comparisons revealed that the third and the eighth posters were consistently rated more inoffensive than all the other posters. These posters were, respectively, highly informative and stressed personal responsibility. Poster 2, which depicted two partly naked males, was rated as being significantly more offensive than all remaining posters. The results appear in Table 7.

The seventh item of the semantic differential required individuals to rate the posters as to how good or bad they were. There were significant differences between mean ratings of the posters ($F=78.19$, $p<.001$). Post hoc comparisons revealed that Poster 3 and Poster 8 were considered significantly better than all the other posters, but there was no significant difference in ratings between these two posters. Poster 3 was highly informative, whereas Poster 8 stressed personal responsibility. Poster 2, which depicted two partly clothed men, however, was considered worse than all of the other posters except for posters 5 and 7. Poster 5 showed Michaelangelo's *David* and clearly relied on humor to convey its message. This may have led to the poster being perceived as fairly innocuous. Poster 7 showed a group of males at the beach. The results appear in Table 8.

The eighth item of the semantic differential required individuals to rate the posters along a continuum as to how moral they were. There were significant differences in the mean ratings of the posters ($F=128.23$, $p<.001$). Post hoc comparisons revealed that the third poster, which was highly informative, and the eighth poster, which stressed personal responsibility, were consistently rated as more moral than the rest. The difference in ratings between these two posters, however, was not significant. Poster 2, which depicted two naked males, was rated significantly more immoral than all the other posters. The results appear in Table 9.

The final item of the semantic differential required individuals to rate the posters along a continuum as to how decent they were. There were significant differences in the mean ratings of the posters ($F=131.16$, $p<.001$). Post hoc comparisons were then performed. The third poster utilizing the informative approach and the eighth poster which stressed personal responsibility were

Table 8
Significant Differences Between Mean Semantic Differential Ratings
Adjective—Good/Bad

Mean	*Poster*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
3.40	1	- -							
4.30	2	*	- -						
1.69	3	*	*	- -					
3.74	4	NS	*	*	- -				
3.76	5	NS	NS	*	NS	- -			
2.88	6	NS	*	*	*	*	- -		
4.07	7	*	NS	*	NS	NS	*	- -	
2.23	8	*	*	NS	*	*	*	*	- -

*pairs of means significantly different at p< .05 level

rated consistently more decent than all of the remaining posters. No significant differences were noted between these two, however. Poster 2 which showed two naked male bodies was rated as being more indecent than all of the remaining posters. The results appear in Table 10.

Discussion

The significant differences noted between semantic differential ratings for different posters suggest that the overall content, message, and design of educational posters has an important bearing on the subjects' reactions to them. The sample generally reacted most positively to two posters: The first was highly informative, highlighting possible myths about AIDS; the second stressed personal responsibility for prevention of the disease. These two posters were significantly and consistently rated as more effective, decent, moral, informative, persuasive, and reassuring than the remaining posters. Conversely, a poster that depicted two naked male chests touching and ad-

Table 9
Significant Differences Between Mean Semantic Differential Ratings
Adjective—Moral/Immoral

Mean	*Poster*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
3.29	1	- -							
4.51	2	*	- -						
1.75	3	*	*	- -					
3.86	4	*	*	*	- -				
3.77	5	NS	*	*	NS	- -			
2.65	6	*	*	*	*	*	- -		
3.93	7	*	*	*	NS	NS	*	- -	
1.82	8	*	*	NS	*	*	*	*	- -

*pairs of means significantly different at p< .05 level

Table 10
Significant Differences Between Mean Semantic Differential Ratings
Adjective—Decent/Indecent

Mean	*Poster*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
3.08	1	- -							
4.45	2	*	- -						
1.63	3	*	*	- -					
3.90	4	*	*	*	- -				
3.86	5	*	*	*	NS	- -			
2.62	6	NS	*	*	*	*	- -		
3.78	7	*	*	*	NS	NS	*	- -	
1.72	8	*	*	NS	*	*	*	*	- -

*pairs of means significantly different at p<.05 level

vocating use of the condom was consistently perceived as less effective and more immoral, shameless, and threatening than remaining posters.

The factors influencing the use of AIDS information are complex (Gerrard & Reis, 1989). Clearly one cannot determine whether these evaluative responses translate into behavior change per se, but the results suggest that the type of message that educational stimuli convey may warrant more careful consideration in planning educational strategies than concern for individual differences within target groups. This is a finding that needs to be investigated further, particularly in light of the fact that comparisons of the means of male and female responses to the posters along all semantic differential levels using Multivariate Analysis of Variance (MANOVA) generally failed to yield statistically significant differences.

Different target groups may also react differently to different posters. In the context of this study it would be reasonable to assume that the sample consisted largely of heterosexuals who appeared to react negatively to stimuli that depicted homosexual lifestyles or activities. As the stimuli became more blatant, so did the strength of the reactions. Yet, as noted, gender differences were not found. This evaluative component appeared to influence perceptions of the posters, most importantly for this study, along a number of dimensions that were identified by the Australian National Advisory Committee on AIDS (Spurgeon, 1988) as barriers to absorption and behavior change. Given that the majority of the sample fell between the ages of 20 to 24, and that they may reasonably be assumed to be sexually active (Health & Welfare Canada, 1989), this sample may prove a particularly useful group to study.

Addressing the uniqueness of individuals' AIDS decisions is an essential component of effective communication (Fischhoff, 1989). Effective education programs may require matching specific target audiences with appropriate educational materials. Yarber (1988) and Stevenson and DeBord (1988) note that factual information is not as important as behavioral and affective education in eliciting behavior change. This may be related to factors such as reactions to stimuli. It may be difficult to distinguish between factual and affective components in education, because the former may be a component of the

latter. In this study the sample clearly appeared to prefer factual material when presented with educational stimuli, but even more apparent was the existence of strong emotional reactions to all the posters. This latter finding warrants further investigation. Most importantly perhaps, the results of this study suggest the messages on educational stimuli may need to be evaluated in field settings before being adopted for widespread use. Semantic differential methodology appears to hold promise as a means of evaluating responses to educational stimuli.

Thus, although much can be learned through attempting to understand the nature and interaction of both micro level and macro level factors in AIDS education, this study emphasizes the need for further research to determine the efficacy of different posters or educational materials on specific individual samples before using them. Once this is done the materials can then be incorporated into comprehensive and multifaceted community-wide or school-based programs in order to provide the most effective primary prevention interventions.

References

Baldwin, J.D., & Baldwin, J.I. (1988a). AIDS information and sexual behavior on a university campus. *Journal of Sex Education and Therapy, 14*(2), 24-28.

Baldwin, J.D., & Baldwin, J.I. (1988b). Factors affecting AIDS-related sexual risk-taking behavior among college students. *Journal of Sex Research, 25*, 181-196.

Fineberg, H.V. (1988a). The social dimensions of AIDS. *Scientific American, 259*(4), pp. 82-89.

Fineberg, H.V. (1988b). Education to prevent AIDS: Prospects and obstacles. *Science, 239*, 592-596.

Fischhoff, B. (1989). Making decisions about AIDS. In V.M. Mays, G.W. Albee, & S.F. Schneider (Eds.), *Primary prevention of AIDS: Psychological approaches* (pp. 168-205). Newbury Park, CA: Sage.

Fisher, W.A. (1980). *Erotophobia-erotophilia and performance in a human sexuality course.* Unpublished manuscript, University of Western Ontario.

Fisher, W.A. (1986). A psychological approach to human sexuality: The Sexual Behavior Sequence. In D. Byrne & K. Kelley (Eds.), *Alternative approaches to the study of sexual behavior* (pp. 131-172). Hillsdale, NJ: Erlbaum.

Fisher, W.A., Byrne, D.E., White, L.A., & Kelley, K. (1988). Erotophobia-erotophilia as a dimension of personality. *Journal of Sex Research, 25*(1), 123-151.

Flora, J.A., & Thoreson, C.E. (1989). Components of a comprehensive strategy for reducing the risk of AIDS in adolescents. In V.M. Mays, G.W. Albee, & S.F. Schneider (Eds.), *Primary prevention of AIDS: Psychological approaches* (pp. 374-389). Newbury Park, CA: Sage.

Gerrard, M., & Reis, T.J. (1989). Retention of contraceptive and AIDS information in the classroom. *Journal of Sex Research, 26*(3), 315-323.

Health & Welfare Canada. (1989). *Canada diseases weekly report, 15*(50), 253.

Kegeles, S.M., Adler, N.M., & Irwin, C.E. (1988). Sexually active adolescents and condoms: Changes over one year in knowledge, attitudes and use. *American Journal of Public Health, 78*, 460-461.

Kelly, J.A., St. Lawrence, J.S., Hood, H.V., & Brasfield, T.L. (1988). Behavioral intervention to reduce AIDS risk activities. *Journal of Consulting and Clinical Psychology, 57*, 160-167.

King, A.J.C., Beazley, R.P., Warren, W., Hankins, C.A., Robertson, A.S., & Radford, J.L. (1988). *Canada youth and AIDS study.* Kingston, ON: Social Program Evaluation Group, Queen's University, Runge Press.

Lenaghan, D.D., & Lenaghan, M.J. (1987). AIDS and education: The front line of prevention. *Futurist, 21*, 17-19.

Maguire, T.M. (1973). Semantic differential methodology for the structuring of attitudes. *American Education Research Journal, 10*(4), 295-306.

Masters, H., Johnson, V.E., & Kolodny, R.C. (1988). *Crisis: Heterosexual behavior in the age of AIDS.* New York: Grove Press.

Norman, C. (1986). $2 billion program urged for AIDS. *Science, 234*, 661-662.

Shernoff, M., & Palacios-Jimenez, L. (1988). AIDS: Prevention is the only vaccine available: An AIDS prevention educational program. In M. Rodway & M. Wright (Eds.), *Sociopsychological aspects of sexually transmitted diseases* (pp. 135-150). New York: Haworth Press.

Sherr, L. (1987). An evaluation of the U.K. government health education campaign on AIDS. *Psychology and Health, 1*(1), 61-72.

Spurgeon, D. (1988). *Understanding AIDS: A Canadian strategy*. Toronto, ON: Key Porter.

Stevenson, M.R., & DeBord, K. (1988). *AIDS awareness: Will knowledge of the facts change behavior?* Paper presented at the meetings of the Society for the Scientific Study of Sex, Chicago.

Strunin, L., & Hingsen, L. (1987). Acquired Immunodeficiency Syndrome and adolescents: Knowledge, beliefs, attitudes, and behaviors. *Pediatrics, 79*, 825-828.

Yarber, W.L. (1988). *AIDS education in schools: Instructional principles*. Paper presented at the meetings of the Society for the Scientific Study of Sex, San Francisco.

Zimbardo, P., & Ebbeson, E. B. (1969). *Influencing attitudes and changing behavior*. Reading, MA: Addison Wesley.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 2, June 1992, 91-103*

James Ryan
The Ontario Institute for Studies in Education

Formal Schooling and Deculturation: Nursing Practice and the Erosion of Native Communication Styles[1]

This study describes one particular example of how modern formal education decultures North American Native students. It documents how instructors in a nursing degree program employ teaching methods that induce Native students to adopt medical practices that are not only contrary to, but that also conflict with their distinctive ways. Native student practice of the scrutinizing conventions that are sanctioned by the medical profession preempt their unique ways of communicating. Furthermore, the wider societal context with which they live encourages them to adopt these ways of interacting in situations outside of medical settings.

Cette étude présente un exemple particulier de la manière dont l'éducation formelle moderne aliène les étudiants Indiens de l'Amérique du Nord. Elle décrit comment les enseignants d'un programme de diplôme d'infirmiers emploient des méthodes d'enseignement qui incitent les étudiants Indiens à adopter des pratiques médicales qui non seulement s'opposent à leurs propres manières, mais aussi sont en conflit avec elles. La pratique par les étudiants Indiens des conventions d'examen qui sont sanctionnées par la profession médicale empiètent sur leurs manières de communiquer. En outre, le contexte social plus large dans lequel ils vivent les encourage à adopter ces modes d'interaction dans des situations en dehors des milieux médicaux.

Educators and social scientists interested in the formal education of Native (and minority) students have been preoccupied over the past two decades with finding ways to improve the success rate of these students.[2] But while the acquisition of skills that accompany success is certainly important to Native people, efforts that are designed exclusively to discover those conditions that impede (or enhance) mastery of these skills overlook other significant by-products of schooling. Indeed, as we have been told, formal education produces effects other than those which are found in official school policies (Apple, 1979; Bowles & Gintis, 1976; Giroux, 1983). This article describes a study of one of these so-called unofficial features of schooling: its role in the deculturation of North American Native people.[3]

Schooling and Deculturation

While use of the term *deculturation* may seem inappropriate to some in this day and age, particularly when employed to characterize Canadian education, it has proven to be a useful concept in analyzing and describing the role of

Jim Ryan is an assistant professor in the Department of Educational Administration. His interests and recent publications are in the areas of Native education, theory and philosophy of education, and educational administration.

education in the spread of European culture throughout Asia, Africa, North and South America in the 17th, 18th, 19th, and even into the 20th century. Carnoy (1974), Altback and Kelly (1978a), and Walsh (1991), for example, recount how Europeans, and later the United States, used Western formal education as part of a planned strategy for extending their domination over the colonies which they were claiming for themselves.[4] According to these authors, these nations saw education as performing two important functions. First, formal schooling was to train the indigenous people of these colonies for roles—invariably subordinate—that suited their colonizers. Second, these schools were to engender in these people an allegiance to a particular and very different political, economic, and moral way of life. Central to these efforts was a move to "civilize" indigenous people, that is, to introduce them to a way of life that resembled that of the Europeans. As Carnoy (1974) maintains, education sought to achieve these ends by degrading "the local 'savages' or 'heathens' into believing that one must be European to be good, strong, able and right" (p. 26). In other words, these colonizing nations used formal schooling as a tool to convince the native people of Asia, Africa, North and South America of the superiority of Western ways and of the advantages that could be accrued from the practice of these ways. A key ingredient then, of education as a form of deculturation was its role in displacing the indigenous culture with another culture by providing the conditions that would propel students to accept and to practice this alien culture.[5]

Europeans, and then Canadians, followed this prescription in their dealings with Canada's Native people.[6] Examples abound of official pedagogical efforts at so-called civilization.[7] Perhaps the most extreme and tragic example is the case of Canada's residential schools.[8] Children bound for these institutions were rounded up, forcibly removed from their families and communities, shipped off to distant schools where they were forbidden to speak their Native tongue, placed on rigid schedules, and taught those skills that their educators felt would permit them to integrate into the dominant culture. While not all Native people were subjected to such treatment, many did attend schools where the curriculum featured exclusively elements of European-based culture, and where teachers adopted tactics implicitly and explicitly geared to convince students of the superiority of this foreign culture and of the advantages of learning skills that had virtually nothing to do with their own culture.

Attitudes toward Native people and Native culture have changed over the past century, however. Residential schools are now a part of history and the integrity of minority (including Native) culture is enshrined in the constitution of the land and in official policy.[9] Despite this fact, however, formal schooling continues to champion Western ideals at the expense of distinctive Native practices that still persist today. Although explicit declarations to this end are not to be found, formal schooling continues in more subtle ways to provide conditions that propel Native students not only to take up Euro-Western ways that are often diametrically opposed to Native practices, but also as a consequence of this acceptance, to forsake elements of their own cultural heritage.

This article outlines a study that documents one particular modern-day instance of this process. More specifically, it describes how the inquisitorial

nature of the communication etiquette required of nurses and nursing students takes precedent over distinctive Native modes of communication that are fundamentally at odds with the former. It shows how elements of a university-based nursing program prompt Native nursing students to adopt and eventually accept ways of communicating that preempt their distinctive ways. Toward this end, I first describe the methods and setting. Next, Native communication etiquette is outlined followed by a portrayal of modern medical communication practice. Finally I describe the means nursing instructors employ to instill these medical conventions in these Native nursing students and the effects these tactics have on these students.

Research Setting and Methods

The study described here focuses on a four-year university-based nursing degree program, and in particular on the 30 Native nursing students who attend this program. These students, all of whom are women save two, vary with respect to age, previous education, work history, and family status, and with the exception of a handful have traveled from communities in the more northerly regions of Canada to attend this on-campus program. All students except one have passed through the one-year preparatory year—now in its fifth year—that was designed to prepare individuals of Native ancestry for the challenges they would face in the degree program.

The research team retrieved data over the course of a university calendar year by scanning various documents and by watching and talking to people in the qualitative research tradition (Bogdan & Biklen, 1982; Woods, 1986). Pertinent documents included course outlines and course and reference material. We also sat in on a number of classes and observed various aspects of classroom interaction. Finally, we conducted in-depth interviews with all Native nursing candidates and instructors. What we looked for and asked about was initially structured by a broad theoretical framework that centered on *support* (Falk & Aitken, 1984; Jeanotte, 1982; Kulig, 1987; Lin, LaCounte, & Eder, 1988), *cultural differences* (Deyhle, 1983, 1986; Dumont, 1972; Greenbaum, 1985; Guilmet, 1981; Philips, 1972, 1983) and *identity* (Cummins, 1986; Dawson, 1988; Luftig, 1983; Ryan, 1989). We initially hoped that information related to these three areas would yield insight into the ways in which these students approached their studies. During the interview sessions, however, we did give instructors and students the latitude to pursue areas they felt were relevant. As the study progressed, and various findings were checked and cross-checked, various themes emerged, including the issue of conflicting communication styles. Once these themes were established, we zeroed in on related issues, searching out documentary sources that would shed light on them, looking for them in the classroom, and asking both students and instructors pointed questions about these topics. Finally, after the all data were collected and written up, I asked various participants to comment on the manuscript. A number of their insightful suggestions were incorporated into this manuscript.

Native Communication Etiquette

Many Native North Americans continue to communicate with one another in ways that differ substantially from the face-to-face interaction commonly

employed by most people of European descent.[10] In order to be properly understood, however, these communication patterns need to be set within the traditional social context in which they were originally practiced. Indeed, conventions associated with communication etiquette were, and to an extent still are, intimately tied to traditional practices of social control, practices that differ markedly from modern Western forms of control. Northern nomadic hunting societies, for example, displayed uniquely egalitarian characters.[11] In other words, these people lived without the formal hierarchical structures that pervade Western culture today. They valued personal autonomy and the right to make decisions about those things that mattered most in their lives. Northern Native people were quite comfortable about others making their own decisions because of the deep abiding trust they had in one another. Ridington (1988), for example, maintains that

> the individualism of Northern Indian people is based on a fundamental social compact, a trust in the individual's social responsibility and informed intelligence. Northern hunting people trust the individual to be responsible, not only to him or herself, but to all other humans and trans-human persons. (p. 107)

The control, then, that one person had over another was decidedly limited within these societies. Efforts to directly influence another or to penetrate in any way another's person space were simply not tolerated.[12]

Although the social context in which many Native people live has changed over the past two centuries, many of these people still retain elements of these traditional values. Philips (1972, 1976, 1983), for example, illustrates how these values continue to influence the communication patterns of the Warm Springs Native people. The issue of control, she claims, plays a central role in the way these people interact with one another. Unlike most Anglos, Warm Springs people who find themselves in the role of listener do not attempt to control the length of the speaker's discourse, nor when they speak, do they strive to set up who will speak next. Indeed, in most cases, the address of the speaker is so general that a response, immediate or otherwise, is not required. People speak up when they are ready. Intrusive actions such as gazing at another, demanding a reply to a query, or asking personal questions constitute inappropriate acts of control and as such are insulting to those to whom these acts are directed, as they violate the fundamental trust that many Native people have in one another. These conventions, however, differ from the communication etiquette sanctioned by the medical profession. Even so, as we shall see, all but a handful of the Native nursing students identified strongly with these Native ways of interacting.

Medical Observational Practice

A central feature of modern medical practice is its preoccupation with what Foucault (1973) refers to as "the gaze," the perpetual and unrelenting scrutiny of patients by medical professionals. At the heart of these pervasive observational practices is an overriding concern to see all, and thereby to know everything there is to know about the patient.[13] While the medical practice of observation is not new, as those familiar with Hippocrates will recognize (Frede, 1990), it took on a central role, according to Foucault (1973), around the end of the 18th century with the reorganization of medical practice and train-

ing. At this time medical professionals adopted more probing, penetrating, and insistent scrutinizing practices that allowed them to diagnose forms of diseases that were now characterized not by essences—as had been the case until this time—but by the visible symptoms that showed themselves in human tissue. But since doctors could not see past the surface skin tissue, they had to figure out what was not immediately visible to the eye. In other words, they had to foresee what dissection would reveal. To accomplish this task they had to employ means other than simple observation, and found themselves engaged in the arts of listening, touching, and questioning, practices that continue to this day. In sum, medical practice revolves around a phenomenon that is totally alien to traditional Native culture, the laying bare of individual men and women before an intrusive and persistent gaze.

Nursing Practice

Nursing practice, as a form of medical practice, also sanctions the art of probing. Indeed collecting information or what both instructional manuals and instructors refer to as "data", constitutes perhaps the most important element of nursing. Griffith and Christensen (1982), for example, maintain that "data collection is the foundation of the nursing process" (p. 29). Observation, as well, is central to this art of data collection, and has been for some time. Florence Nightingale (in Bradley & Edinberg, 1982) is reported to have told her nursing colleagues that "if you cannot get into the habit of observation one way or another, you had better give up being a nurse" (p. 29). In a modern-day appraisal of observational techniques Bradley and Edinberg (1982) contend that "Observation is the sine qua non of nursing. Nurses primarily obtain information about clients in two ways: they *listen* to what clients tell them, and they *observe* objective signs and symptoms" (p. 29).

Because of the importance that the nursing profession accords to the acquisition of information, those who teach prospective nurses give priority to equipping these people with communication skills that will allow them to perform this task. In fact, most nursing instructors in this particular program believe that appropriate communication skills are essential to nursing. Marion, a veteran instructor, maintains that "if you can't communicate so that people can understand, errors can be made, people can die." Margaret, another instructor, outlines the vital connection between communication, information gathering and the helping relationship.

> *If we can't communicate then we can't gather data. If you can't gather data, you can't figure out how you can help. Therefore all the skills in the world aren't going to help if you don't know what's the matter.*

Appropriate communication, at least as it pertains to nursing, requires that the nurse establish a relationship with the patient that will induce him or her to reveal all. One particular technique favored by nursing manuals and instructors involves the projection of "warmth behaviors." This method has proven to be attractive to many nurses because, as Smith (1986) contends, "clients who sense your warmth are more likely to engage in dialogue with you and provide you with the information about their health condition you need to make a nursing diagnosis" (pp. 48-49). Nursing instructors attempt to get their charges to become proficient in projecting those warmth behaviors

that would lead patients to believe that the attending nurse is a warm and caring person, and that it would be in the patient's best interests to confide in this nurse. Madeline, a nursing instructor, elaborates on those traits that she believes project a caring image.

> *We're talking about people whose face indicates that we really like talking to you. We're smiling, eyes that look interested, eye contact, open posture where we're indicating that "Yes, I think you're really important, and I'm willing to listen to you." So we use a lot of this warmth behavior. We try and encourage respect by calling the person by his or her name, giving them the time, giving them the response when they're asked. We look at things like empathy, being able to put yourself in that person's position, if only for a moment, and share with them that you understand, but that requires a response. We talk about nonverbal skills that indicate warmth, like holding someone's hand.*

In sum, the nursing profession requires that those who practice its tenets take an inquisitor's stance toward patients. Those who employ these methods believe that a relationship of this sort will allow the nurse to wrest information from the patient that will eventually assist the medical professional in making a prognosis.

Instructing Nursing Students

Nursing instructors in this program employ a number of instructional techniques to ensure that student nurses acquire the attitudes and skills necessary to extract appropriate information from patients. They address these skills both in seminar sessions where up to six students are in attendance, and in actual health care facilities. In most cases these instructors review the skills in the seminar session, and students attempt to put what they have learned into practice in placement settings. Students generally participate in various exercises and complete assignments that are designed to sharpen their information gathering skills. One assignment requires students to tape an interview session with a patient, transcribe it, analyze it, and hand it in to their instructor, who in turn offers suggestions. In the seminar session, instructors may also plan activities that depart from the customary discussion or lecture format. One popular exercise involves a series of audiotapes. Activities associated with these tapes include, for example, having students (a) listen to a fictional patient talk, (b) write down what they believe to be the "correct" meaning of these utterances, (c) and then discuss with classmates and instructors strategies for discerning these correct meanings. Another popular strategy is role playing. Instructors introduce students to contrived situations that require one student, who plays the role of the nurse, to pursue a line of questioning that will lead him or her to elicit certain information from the other actor, who plays the role of the patient. Comments and advice are generally offered both by participants and observers after the performance. Remarks from one such session included, for example, such statements as "it was inappropriate to talk loud," "I didn't use open-ended questions," "I felt I was firing questions," "be accepting about receiving information," "remember it is important to listen," and a comment that most Native students would identify with, "I need more help in probing."

This last observation by one of the Native students reflects the difficulty that many of these students have adopting conventional nursing communica-

tion patterns. Indeed, most Native students are very uncomfortable, at least initially, as they are pushed to take an inquisitor's stance. Janice, a Native student, tells us, "I had a hard time at first because I'm Native and pretty shy by nature." Jill, on the other hand, draws a connection between her cultural heritage and the difficulty she has doing some of the things that are required of her by her instructors.

> *The Native people are very shy. I guess we are brought up like that, and there are even different ways of body posture. Like we're not allowed to look at someone's eyes, for certain periods of time. The nurse will say, "Okay, when you're talking to someone, look at their face." But for us that's not our belief. We always have to put our heads down or bow our heads. That's what I was taught. It's a behavior that was taught when we were very small and is part of our culture. To some people (non-Natives), it's as if we're ignoring them. But to us, it's out of respect.*

All instructors are aware that most Native students have problems employing these practices. They cite their lack of aggression and assertiveness, traits that show up, for example, in their unwillingness to look into patients' eyes or to ask direct questions. As one Native student told an instructor, "I wouldn't ask those kinds of questions of my grandmother, so why would I ask them of somebody I don't even know?" Mary Ann is only one of many instructors who has noticed that Native students "feel a little reluctant to go in and ask the appropriate questions, and therefore to get the appropriate data base that's required." Another instructor says that "they're not assertive. If the client is busy or the curtain [around the bed] is closed, they're very careful about not asserting themselves." Although most instructors recognize and appreciate the unique ways in which many Native students interact with others, they nevertheless feel that it is their obligation to see that they master conventional nursing communication techniques and forgo elements of their cultural heritage. Maureen, perhaps more than most instructors, realizes what is at stake here. She states that

> *The expectation here is that they will communicate as we do. I have to say that they have to learn that they cannot communicate [in their distinctive ways].... I think I can appreciate the way Native people communicate, but I have to teach them to communicate [as a] generalist. So they have to change.*

Because many Native students have difficulty stepping into the role of inquisitor, instructors find they must take extra or unusual measures to ensure that these students do not employ their distinctive interactive practices either in the classroom or in hospital wards. Instructors may feel a need to intervene at times. This happened in one particular situation where two Native students were involved in a role play. This role play, however, took on a different character than the previous interchanges. Unlike the exchanges that had already taken place, the two women, playing the respective roles of nurse and patient, did not launch into an immediate dialogue. Instead both of them sat facing each other without uttering a word. Time went on, and still nothing was said. They eventually began to display their discomfort, twisting in their chairs, occasionally glancing at the walls, ceiling, and floor, and generally avoiding the eyes of the other. After a full two minutes had passed, the instructor intervened, exhorting the women to speak. Finally, after another

pause of a shorter duration, they began. After the class the women approached the instructor and told her that they were reluctant to begin not because they were uneasy about performing in front of the class, but because they felt it was inappropriate for either of them to invade one another's space by initiating an intrusive dialogue. For the instructor their silence was unacceptable, and she felt she had an obligation to put an end to it, regardless of the fact that she personally had a great deal of respect for Native culture. In the end she felt that these practices were simply out of place in this particular setting.

Other instructors prefer other tactics in dealing with students' reluctance in these matters. Margaret, for example, prefers to set up a role play situation in which just she and the student are present. She feels that some students may be more willing to practice when there are no spectators. Mildred, on the other hand, insists on asking reluctant students rhetorical questions to rationalize their participation in probing efforts. She says "I ask them why they are here. I then answer that we are here to learn and that they should try and work around the lack of assertiveness." Finally, Mona feels that acknowledging their feelings will assist students in overcoming their discomfort. Asked what she says to these students, she replied,

> *I say something like, "How do you feel when you are in a room with a patient and you are getting ready to ask questions? What are you feeling?" ... I think recognizing feelings is a very important part in overcoming these obstacles.*

Inevitably, however, most students overcome their reluctance and embrace this role, at least to a point that satisfies their instructors.[14] Indeed, they have little choice in the matter: they must conform or risk failure. Conforming, however, is not an easy matter. June, for example, says that "when we entered the program the coordinator told us we had to mesh the Native and non-Native ways. It's really hard to." Furthermore, being placed in an environment that favors social practices that are diametrically opposed to those valued by one's own cultural heritage can prove troubling to students in other ways. Jillian believes that "the teachers think there's something wrong with a Native student because they don't talk, or they don't communicate [in ways that teachers expect]."

Discussion

Although popular opinion may support the notion that Western societies have progressed a long way from the days when Europeans forcibly imposed their culture on peoples from other parts of the world, the deculturation process that was an integral part of these colonial ventures persists today. It continues to show itself in many forms, including educational institutions that cater to students whose cultural heritage differs from those of Euro-Western descent. Today, however, it works in more subtle ways. But although it may not be part of a planned strategy, modern-day deculturation nevertheless works to replace one set of cultural practices with another, extending the domination of the latter over the former through schooling practices that both implicitly and explicitly convey the supposed superiority of Western ideals. This process negates not only the culture of these subordinate groups, but also their history and sense of self.

Modern-day deculturation is not a uniform or determinant process. It does not, to cite a familiar phrase, work itself out behind the backs of men and women. Rather, intelligent human beings are part of this process: they make choices day in and day out that produce and reproduce these dominant relations. The case described here illustrates just one instance of how deculturation works itself out on a daily basis in an educational setting. As we have seen, different instructors employ different means to equip Native nursing students with the skills they believe are essential to the profession. But although various instructors may employ a wide variety of tactics, they all have one thing in common: an unquestioned allegiance to, and belief in, the inquisitorial method of attending to patients. The superiority of this exclusively Western practice is a fact that is taken for granted by all instructors, and they take great pains to ensure that Native students develop the skills associated with it, even though its employment negates the Native students' own cultural practices. Even those instructors who have an understanding and appreciation of Native communication etiquette will press the issue with students because apart from their commitment to these practices, they believe that acquiring such skills will inevitably be in the students' best interest.

On the other hand, student acceptance of inquisitional relationships with patients is neither automatic nor uniform. Many students resist, at least initially, those tactics instructors use to get them to accept and to adopt these practices. For many Native students, especially those who are close to their historical roots, deciding to employ these skills is no easy matter, for learning them involves much more than, say, learning to ride a bicycle. This process is emotionally charged. Making up one's mind to look someone straight in the eye for sustained periods of time, or asking a personal question of a patient means that a Native student must in some way acknowledge the superiority of these practices over Native ways of communicating. This confrontation of cultures is much more direct than, for example, would be case of simply not teaching traditional Native healing practices, and as a result can be more painful for a Native person whose sense of self is ultimately tied to this way of life. Indeed the negation of these distinctive practices is accompanied by a negation of the self of each of those Native students who feel at least some attachment to their heritage.[15]

But is this particular instance an isolated event? In other words, one might ask, wouldn't Native students simply do what they had to do in the classroom and the health care setting to satisfy instructors or their professional duty, and then revert to their distinctive ways of communicating in other settings? To answer this question, we must acknowledge that the dismantling and replacement of Native ways with Euro-Western practices does not happen in isolation. The imposition of inquisitorial relationships goes on side by side with the imposition of other dismantling Euro-Western practices, both in educational institutions and out, which both reinforce and build upon one another. The individualism that pervades Western institutions, including education, provides a case in point. These individualizing practices work with other related practices to break down the distinctive Native communal orientation to life (Ryan, in press). They also constitute but one of many sorts of practices that provide conditions that prompt Native students to employ inquisitorial

techniques in their interaction outside of these institutions. Indeed these questioning techniques come in handy in a society where men and women are dealt with, and are forced to deal with others, on an individual (as opposed to group) basis.[16] In such circumstances one must be prepared to step forward to both elicit and provide information to a wide variety of individuals on whom one depends for the essentials of life. Unlike traditional Native times when people had greater control over their own destinies and were considered first and foremost a part of a group, living in the dominant Western culture now requires that they elicit and supply information on an immediate basis to such people as doctors, teachers, colleagues, bosses, shopkeepers, lawyers, civil servants, repair people, bureaucrats, social workers, and police officers, among many others. Although at times learning the rudiments of this communication style may be a difficult process, these ways of relating to others nevertheless help Native nurses to get along in a world that requires that they engage in these sorts of interactions. As such, most are quite willing to make use of these new skills in contexts other than nursing situations.

Conclusion

Many Native people today find themselves in an uncomfortable situation. On one hand, many look on with disfavor as they see their cultural heritage slowly slipping into the past, a consequence of, first, the depletion of game animals, among other resources, that at one time sustained their ancestors, and, second, the overwhelming nature of Euro-Western forms of life that are slowly but surely dismantling what is left of their relationship to the past. On the other hand, Native people are put in the awkward position of being compelled to contribute to this process of deculturation. Indeed, doing the things they must do to survive in this imposed environment requires that they forsake not only the more visible elements of their past such as clothing or hunting rituals, but other fundamental conventions of their heritage that are not as evident to non-Natives, such as ways of interacting with one another. Formal schooling typifies the contradictory nature of this process. While formal education may provide Native students with skills that help them get along in the dominant society, these same skills also contribute to the erosion of distinctive Native practices.

What, then, is to be done about this state of affairs? Do Native people have no other recourse than to stand powerlessly by and watch their cultural heritage slip into the past? If alternatives to current practices are to be found, they will hinge on Native control of Native institutions, including schools. Only when Native people have the right to decide, without intrusive government intervention, how they will live their lives, will they be able to make substantive and meaningful changes to current institutional practice. Control of this nature will enable them to critically analyze, evaluate, and appropriate elements of current practice that will best serve their needs. Such revered institutions as medical science, or individualism and related practices that surface in schools, for example, need not escape intense scrutiny.[17] Such scrutiny will in turn assist Native people in deciding which established conventions are worth keeping and which are best discarded. There is an urgency in this task, however, for if not enacted in the near future, there will come a day when Western thought and practice has become so ingrained in Native

people that many may not have the will to do what has to be done to arrest, or at least alter, the process of deculturation.

Notes

1. Funding for this study was provided by Indian and Northern Affairs Canada (INAC).
2. Typical of these efforts is the ongoing dialogue in *Anthropology and Education Quarterly* that has included Cazden (1983), Ogbu (1987), Erickson (1987), Trueba (1988), and Foley (1991), among many others. The debate has centered on whether interactive or structural factors, or a combination of both, constitute the more powerful explanation for minority student failure and success.
3. It is recognized here that although the majority of Native people in Canada do not practice their traditional culture as their ancestors once did, a consequence in part of the deculturation process that began 500 years ago, many of them still retain elements of this heritage despite the fact that such practices might not be all that evident to non-Native observers. As Erickson and Mohatt (1982, p. 166) contend, it is "precisely because of the presence in everyday life of modern dress, cars, outboard motors, television sets, and fluency in English that these relatively subtle aspects are likely to go unrecognized." And while these less obvious ways of life may well not be "original" or "authentic" dimensions of Native culture, they remain distinctly Native in character (see, e.g., Ryan, in press).
4. Walsh (1991) gives us an illuminating look at how the United States attempted to use education to peacefully colonize Puerto Rico, a prize it had acquired in the Spanish-American war.
5. Altback and Kelly (1978b) maintain that Europeans never really did intend that colonial schools would enable the colonized people to integrate into their society. What integration there would be was intended to be marginal and to serve the needs of the colonial power.
6. The situation of those Native people living in Canada changed from one of classical to internal colonialism (Altback & Kelly, 1978b) as the descendants of Europeans (i.e., Canadians) took control of the land.
7. An account of early efforts at educating Native people can be found in Thwaites (1906) and Jaenen (1986). Iverson (1978) provides a sampling of pedagogical efforts to civilize American Native people over a 350-year period. Finally, for an account of similar efforts in the 1950s on the part of Canadian government officials see Ryan (1988). These efforts, as have most other efforts of this sort, have met with little success from the vantage points of both Native and non-Native parties.
8. See, for example, Persson (1986).
9. Mock and Masemann (1990), for example, found that 69 of 129 Ontario school districts either have ethnocultural equity policies or are in the process of developing them. The Canadian Charter of Rights and Freedoms states that all individuals are "equal before the law without discrimination based on race, national or ethnic origin, colour, religion, sex, age or mental or physical disability."
10. I do not mean to imply here that all North American Native people are the same, a belief promulgated by the colonial European powers, among others. North America is populated by a wide range of aboriginal people who differ in many respects. Some common themes, however, transcend tribal boundaries. This was the case with the notion of social control and communication etiquette. Both northern hunting societies and the more southerly Warm Springs people, for example, displayed marked similarities with respect to these conventions. Furthermore, I felt justified in using these examples because many of the Native nursing students I interviewed identified with the northern people and their ways of communicating.
11. See, for example, Henrikson (1973) and Mailhot and Michand (1965).
12. Ridington (1986, p. 105) draws a distinction between the egalitarian societies of these northern people and the Western system of social hierarchy. He says that "in such an egalitarian society the authority of individual judgments requires respect for individual autonomy and nonintervention, while in a system of social hierarchy individualism is viewed as the mechanism by which some people gain the power to intervene in the lives of others."
13. The relationship of patient and doctor differs considerably from the relationship that many Native people once had with one another. In modern medical practice the medical professional assumes virtual control over the patient, a situation that would not have been possible in many Native societies.

14. One student who was not entirely comfortable with asking point-blank questions would, if at all possible, wait for the patient to tell her what she was there to find out. She said, "You have to approach them indirectly, so that they tell you what they have.... You let them speak."
15. Cummins (1986) draws a relationship between identity and the success of minority students. He maintains that young people who are not positively oriented to their own culture or to the dominant culture, who perceive themselves as inferior, and who are alienated from their own cultural values generally do not succeed at the process of formal schooling. For other statements on the impact of formal schooling on identity and self-image, see Luftig (1983), Dawson (1988), and Ryan (1989).
16. For a description of the origins of the individualizing practices of modern-day society, see Foucault (1979).
17. Ehrenreich (1978), for example, contends that factors other than personal medical care have been responsible for the reduction in morbidity rates. According to him, better standards of living that include housing and nutrition and selected environmental factors such as appropriate water supplies and sanitary services have played a more important role in the improvement in health levels over the past century. But while most Native people live longer today than many of their ancestors, the state of their mental health has deteriorated at an alarming rate. To cite an example typical of many northern areas, 43 young people in 28 isolated northern Ontario communities have taken their own lives over the past five years (Sanders, 1991).

References

Altback, P., & Kelly, G. (Eds.). (1978a). *Education and colonialism*. New York: Longman.

Altback, P., & Kelly, G. (1978b). Introduction. In P. Altback & G. Kelly (Eds.), *Education and colonialism* (pp. 1-44). New York: Longman.

Apple, M. (1979). *Ideology and curriculum*. London: Routledge and Kegan Paul.

Bogdan, R.C., & Biklen, S.K. (1982). *Qualitative research for education: An introduction to theory and methods*. Boston: Allyn and Bacon.

Bowles, S., & Gintis, H. (1976). *Schooling in capitalist America: Educational reform and the contradiction of economic life*. New York: Basic.

Bradley, J.C., & Edinberg, M.A. (1982). *Communication in the nursing context*. East Norwalk, CT: Appleton-Century-Crofts.

Carnoy, M. (1974). *Education as cultural imperialism*. New York: McKay.

Cazden, C. (1983). Can ethnographic research go beyond the status quo? *Anthropology and Education Quarterly, 14*, 33-41.

Cummins, J. (1986). Empowering minority students: A framework for intervention. *Harvard Educational Review, 56*(1), 18-36.

Dawson, J. (1988). If my children are proud: Native education and the problem of self-esteem. *Canadian Journal of Native Education, 15*(1), 43-50.

Deyhle, D. (1983). Between games and failure: A micro-ethnographic study of Navajo and testing. *Curriculum Inquiry, 13*(4), 347-376.

Deyhle, D. (1986). Success and failure: A micro-ethnographic comparison of Navajo and Anglo students' perceptions of testing. *Curriculum Inquiry, 16*(4), 365-389.

Dumont, R.V. (1972). Learning English and how to be silent: Studies in Sioux and Cherokee classrooms of Indian students. In C.B. Cazden, V.P. John, D. Hymes (Eds.), *Functions of language in the classroom*. New York: Teachers College Press.

Ehrenreich, J. (1978). Introduction: The cultural crisis of modern medicine. In J. Ehrenreich (Ed.), *The cultural crisis of modern medicine*. New York: Monthly Review Press.

Erickson, F. (1987). Transformation and school success: The politics and culture of educational achievement. *Anthropology and Education Quarterly, 18*(1), 335-356.

Erickson, F., & Mohatt, G. (1982). Cultural organization of participant structures in two classrooms of Indian students. In G. Spindler (Ed.), *Doing the ethnography of schooling* (pp. 132-174). Toronto, ON: Holt, Rinehart & Winston.

Falk, D., & Aitken, L. (1984). Promoting retention among American Indian college students. *Journal of American Indian Education, 23*(2) 24-31.

Foley D.E. (1991). Reconsidering anthropological explanations of ethnic school failure. *Anthropology and Education Quarterly, 22*(1), 60-86.

Foucault, M. (1973). *The birth of the clinic: An archaeology of medical perception*. New York: Vintage.

Foucault, M. (1979). *Discipline and punish: The birth of the prison*. New York: Vintage.

Frede, M. (1990). An empiricist view of knowledge: Memorism. In S. Everson (Ed.), *Companions to ancient thought 1: Epistomology*. New York: Cambridge University Press.
Giroux, H. (1983). *Theory and resistance in education: A pedagogy for the opposition*. South Hadley, MA: Bergin & Garvey.
Greenbaum, P.E. (1985). Nonverbal differences in communicative style between American Indian and Anglo elementary classrooms. *American Education Research Journal, 22*, 16-33.
Griffith, J.W., & Christensen, P.J. (1982). *Nursing process: Application of theories, frameworks, and models*. St. Louis, MO: Mosby.
Guilmet, G.M. (1981). Oral-linguistic and nonoral-visual styles of attending: Navajo and Caucasian children compared in an urban classroom and on an urban playground. *Human Organization, 40*, 145-150.
Henrikson, G. (1973). *Hunters in the barren: The Naskapi on the edge of the white man's world*. St. John's, NF: Memorial University Institute of Social and Economic Research.
Iverson, K. (1978). Civilization and assimilation in the colonialized schooling of Native Americans. In P. Altback & G. Kelly (Eds.), *Education and colonialism*. New York: Longman.
Jaenen, C.J. (1986). Education for francization: The case of New France in the seventeenth century. In J. Barman, Y. Hebert, & D. McCaskill (Eds.), *Indian education in Canada: Volume 1 The legacy. (pp. 45-63)* Vancouver, BC: University of British Columbia Press.
Jeanotte, L. (1982). *A study of the contributing factors relating to why American Indian students drop out or graduate from educational programs at the University of North Dakota*. Paper presented at the Annual Meeting of the American Educational Research Associates, New York.
Kulig, J. (1987). Native American registered nurses and state board examinations. *Journal of American Indian Education, 26*(2), 29-38.
Lin, R., LaCounte, D., & Eder, J. (1988). A study of Native American students in a predominantly white college. *Journal of American Indian Education, 27*(3), 8-15.
Luftig, R.L. (1983). The effects of schooling on the self-concept of Native students. *The School Counselor, 30*(4), 251-260.
Mailhot, J., & Michand, A. (1965). *North West River: Etude ethnographic*. Quebec, PQ: Centre d'Etudes Nordiques.
Mock, K., & Masemann, V. (1990). *Implementing race and ethnocultural equity policy in Ontario school boards*. Toronto, ON: Queen's Printer.
Ogbu, J.U. (1987). Variability in minority school performance: A problem in search of an explanation. *Anthropology and Education Quarterly, 18*(1), 312-334.
Persson, D. (1986). The changing experience of Indian residential schooling: Blue Quills, 1931-1970. In J. Barman, Y. Hebert, & D. McCaskill (Eds.), *Indian education in Canada: Volume 1 The legacy. (pp. 150-168)* Vancouver, BC: University of British Columbia Press.
Philips, S. (1976). Some sources of cultural variability in the regulation of talk. *Language in Society, 5*, 81-95.
Philips, S.U. (1972). Participant structures and communicative competence: Warm Springs children in community and classroom. In C.B. Cazden, V.P. John, D. Hymes (Eds.), *Functions of language in the classroom* (pp. 370-394). New York: Teachers College Press.
Philips, S.U. (1983). *The invisible culture: Communication in the classroom and community on the Warm Springs reservation*. New York: Longman.
Ridington, R. (1988). Knowledge and power in the subarctic. *American Anthropologist, 90*(1), 98-110.
Ryan, J.J. (1988). Economic development and Innu settlement: The establishment of Sheshatshit. *Canadian Journal of Native Studies, 8*(1), 1-25.
Ryan, J.J. (1989). Disciplining the Innut: Normalization, characterization, and schooling. *Curriculum Inquiry, 19*(4), 379-403.
Ryan, J.J. (in press). Eroding Innu cultural tradition: Communality and individualization. *Journal of Canadian Studies*.
Sanders, C. (1991, May 14). Nishnawbe-Aski comes to grip with high suicide rate among young Natives. *Chronicle-Journal*, Thunder Bay: Thompson Newspapers.
Smith, S. (1986). *Communications in nursing*. Toronto: McGraw-Hill Ryerson.
Thwaites, R.J. (Ed.). (1906). *The Jesuit relations and allied documents*. 71 vols. Cleveland, OH: Borrow Brothers.
Trueba, H. (1988). Culturally based explanations of minority students' academic achievement. *Anthropology and Education Quarterly, 19*, 270-287.
Walsh, C.E. (1991). *Pedagogy and the struggle for voice: Issues of language, power, and schooling for Puerto Ricans*. Toronto, ON: OISE Press.
Woods, P. (1986) *Inside schools: Ethnography in educational research*. New York: Routledge & Kegan Paul.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 2, June 1992, 105-119*

Margaret Haughey
and
L. MacElwain
University of Alberta

Principals as Instructional Supervisors

Since recent literature on excellence in schools stressed the positive effect that principals can have on quality instruction, instructional supervision has received renewed emphasis in writings on school administration (Owens, 1991). Principals are expected to empower their staffs, to enhance commitment to student learning, to spend a major part of their day improving the instructional program, and to expand opportunities for teacher growth. At the same time principals are required to evaluate teachers and monitor programs. As Pajak (1990) has noted, "Instructional supervision, clinical supervision and classroom observations of teachers by administrators are virtually synonymous in the minds of practitioners" (p. 78). The study reported in this article explored eight principals' understandings of instructional supervision.

Puisque la littérature courante concernant l'excellence dans les écoles a mis l'accent sur l'effet positif que les directeurs peuvent exercer sur la qualité de l'enseignement, la supervision pédagogique a connu un regain d'intérêt dans les écrits sur l'administration scolaire. Les directeurs doivent donner du pouvoir à leur personnel, accroître la responsabilité envers l'apprentissage des élèves, passer la majeure partie de leurs journées à améliorer le programme d'enseignement, et augmenter les occasions propices au développement professionnel des enseignants. En même temps, les directeurs ont l'obligation d'évaluer les enseignants et de contrôler les programmes. L'étude présentée dans cet article a exploré les vues de huit directeurs au sujet de la supervision pédagogique.

Instructional Supervision

Interest in the topic of instructional supervision in education has inspired considerable research and resultant literature. Oliva (1984) identified three themes prevalent in the writings of specialists on supervision: (a) a focus on the behavior of supervisors, (b) in assisting teachers, (c) for the ultimate benefit of the student. There is general agreement among writers that the focus of instructional supervision is the enhancement of student learning. Where writers differ is in the ways they suggest instructional supervision should be practiced. Some writers emphasize the role of the supervisor. For example, Wiles and Bondi (1986) defined supervision as "a general leadership function that coordinates and manages those school activities concerned with learning" (p. 10). Smith and Andrews (1989) identify four roles that they see as essential aspects of an instructional leader: resource provider, instructional resource communicator, and visible presence. Other writers (e.g., Glickman, 1985;

Margaret Haughey is an associate professor in the Department of Educational Administration. Les MacElwain is a principal in the Edmonton Public School District.

Smyth, 1988) stress the importance of the involvement of the teacher in the process. In a recent work Sergiovanni (1987), building on the work of Costa and Garmston (1985), presented a reflective practice model of supervision in which he proposed that since teachers vary in their goals and learning styles, supervisors should be responsive to these differences in the ways they supervise.

> At the heart of this view of supervision is the notion that teachers have work goals that are important to them. Given the opportunity, they will work very hard at achieving them. The nature of these goals is influenced by growth stages, cognitive complexity levels, learning styles, and motives that teachers bring to school. Supervisory options and styles should respond to these differences among teachers, for such responsiveness makes it easier for work goals to be realized. (Sergiovanni, 1987, p. 225)

Yet, despite the general agreement among writers about a definition of instructional supervision, researchers have found that principals vary in the ways they describe their work as instructional leaders.

Through data gathered in a year-long study of several outstanding principals (Dwyer et al., 1985), Dwyer (1986b) learned "that there is no single image or simple formula for successful instructional leadership" (p. 15). Instead, each situation had to be understood in the context of the community and the aspirations of the principal.

> The studies allowed a very different concept of leadership behavior of school principals to emerge. This concept is one that visualizes instructional leadership accruing from the repetition of routine and mundane acts performed in accord with principals' overarching perspectives on schooling. (Dwyer, 1986a, pp. 66-67)

McEvoy (1987), who examined the data from Dwyer's study to identify the ways principals exercised instructional leadership, stressed the importance of everyday acts of principals, especially informal talk, in promoting staff development. In comparing what was written about instructional leadership with findings from a number of studies, Morris, Crowson, Porter-Gehrie, and Hurwitz (1984) observed that "studies on what administrators actually do have discovered that principals spend relatively little time directly attending to teaching and learning" (p. 74) but concluded, "the principal's influence, though, may be more indirect, creating a favorable climate for learning: the most effective role may be supportive rather than supervisory or evaluative" (p. 74). This was confirmed in their own study and described in a section titled "Instructional leadership by indirection" (p. 94). Although these principals were not heavily involved in classroom observation and teacher evaluation procedures, they considered themselves to be instructional supervisors. Like the principals in Dwyer et al.'s (1985) study they had little difficulty in describing "their efforts to shape and form the content of instruction and color the ambience of their schools" (p. 4).

Research Methodology

Dwyer's (1986b) research provided a starting point for this study. The major research questions were: How do principals understand the notion of instructional supervision, and how do they put their understanding into practice?

The theoretical perspective of the study is based on a phenomenological understanding of interactions where the meanings that participants give to their actions are crucial in providing an interpretation of their activities. In this study, eight male principals from one large urban school system were interviewed in order to discover what these principals as instructional supervisors do. Although these principals were not considered representative of principals in their school district, they were chosen to include a variety of years of administrative experience and level of school. Each principal had been identified by a number of colleagues and senior administrators as being someone who had a positive reputation in working with teachers. One was a high school principal, three were principals of elementary-junior high schools and four were elementary school principals. While two had less than three years of administrative experience, the remainder had been principals for at least five years including two who had been administrators for over 25 years. When approached, all the identified principals were excited about being involved in the study, so all eight were included.

Data gathering strategies included nonparticipant observation and reflective interviews. Each principal was visited on between two and four occasions. The first half-day visit began with an observation period during which the researcher shadowed the principal and took brief notes as the principal carried on with the work of the school. There were two purposes for the observation. It provided a context for the interviews since the researcher and principal were able to refer to actual activities that grounded the discussion of instructional supervision. The observation period also helped the researcher become aware of the culture of the school. On the next visit to the school, the principal participated in a taped reflective interview. This interview was structured to explore the two research questions on the meaning of instructional supervision and the activities of an instructional supervisor. Within that format the principal provided the direction for the interview while the researcher listened, probed where appropriate, and tried to obtain concrete examples of the principal's actions. On average, interviews lasted at least two hours. These interviews were transcribed and returned to the participants for verification. The researcher met with each principal to ensure that the transcript was an accurate reflection of the conversation and there was nothing that, in retrospect, should be changed or discarded. During this taped conversation, the researcher also asked questions and sought clarification about statements in the transcripts.

Intensive readings of the transcripts led to the categorization of the data. Each principal's transcripts were read as a unit but considered separately from those of other respondents. The transcripts were analyzed in two ways. First, as the transcript was read, a category name was written in the margin beside each section where the topic was discussed. These sections could be a phrase, a sentence or a paragraph depending on the content. Then a second more detailed reading, line by line, and word by word, helped explore and identify more substantial topics, which were also labeled and written in the margins. When each respondent's transcripts had been analyzed, the category labels from all the transcripts were examined for similarities, and like topics were grouped together. This process provides a way of grouping the responses

while allowing for the diversity of responses under each topic. Once the data were written and returned to the principals, the researcher met with them and asked whether their understandings had been accurately captured in the description of the findings.

The Actions of Supervisors

Many of the principals described what they did as part of their response to what instructional supervision meant for them. They replied to the question about their understanding of instructional supervision with specific examples of what they did or tried to do. For them, their actions as instructional supervisors were both illustrative of and embodied their meanings of instructional supervision. Some had difficulty separating the definition from the action. As one said in a tone of frustration, "Instructional supervision? It's everything I do!" This entanglement of purpose and action became one characteristic of these principals' understanding of instructional supervision. The principals described numerous ways in which they acted as instructional supervisors, and these responses have initially been organized according to four broad categories: setting and maintaining expectations; providing resources and support; observing others and being observed; and encouraging the involvement of others. Within each category, the diversity of responses by the principals is displayed.

Setting and Maintaining Expectations

For many of the principals this activity started at the beginning of the year or when they arrived at a school as its new principal. They felt that the development of a school philosophy or vision, either alone or in concert with staff, was an essential foundation from which the instructional activities of the school would evolve. For example, one principal recalled that for him:

> *The opening two days of school are probably the most important two days of the whole year. And so you plan to the nth degree for those two days to make sure that the staff realizes, "Hey, this is our focus. This is where we're going, this is our vision for the school, these are our priorities." You know, our objectives are set from those and then, I think people understand.*

Two principals talked about being new to their schools and reinforced the importance of establishing "mission statements" as a first task. As one pointed out:

> *We began as a staff, right from the very beginning, to adopt a statement of beliefs. I did not even want to start off the school year without the adoption of a statement of beliefs, in that I believe that schools must be student-centered. In sharing that with the staff, what does that mean? And so that we begin the school year with that focus, so that at any time during the year, if there is a problem or concern, we can address it in terms of meeting student needs and in that instruction involves all areas and aspects of student involvement, then in which way can we make the instructional component supportive to student needs?*

A new elementary school principal stated, in regard to his first year, "and so we had to spend a lot of time to come up with some sort of vision statements. What is our curriculum supposed to do for our kids?"

The acceptance of the mission statement by staff was seen by principals as being crucial to their effectiveness as instructional supervisors. The willingness of the teachers to "buy into the mission of the school" was viewed by many principals as being of utmost importance. The adoption of a single school philosophy provided a common focus for all staff members and allowed them to work together in a supportive manner. One principal concluded:

> *Well, things that help you be an effective supervisor, naturally, are that the staff that adopts the philosophy or mission of the school or indeed a philosophy or mission similar to the one that you've talked about. In other words, making kids feel good about themselves so that you're offering the best curriculum and the best delivery that you can with what you have, and if everybody's headed in that direction, then your job is 10 times easier, absolutely. You have to work with good people. You can't work in a vacuum. So the staffing component is ultra, ultra important.*

Principals saw this as being much preferable to a situation where there was not a shared mission as these situations had great potential for conflict and non-cooperation.

As well principals viewed healthy principal-teacher relationships as being invaluable to the enhancement of instructional supervision. One principal stressed the importance of teachers' attitudes toward the principal and saw great benefits result from "the general staff willingness to have me in the classroom." Another respondent described the difficulties resulting from poor communication between principals and their staff:

> *If the communication lines are not open because you're new in the school or because of lack of trust, then information that should be coming to you doesn't come to you. Or it comes to you at a time when things are already very bad. You see, you're constrained by your own access to information, for one thing.*

Recognizing the importance of good teacher-principal relationships, principals worked hard to establish these relationships. The importance of interpersonal relations was stressed by one principal whose strategies were "supportive listening, by remaining visible and approachable, by attempting to break down barriers and by letting them know that you have feet of clay."

Another form of clarifying goals and expectations occurred at an individual level. Three of the principals mentioned having held conferences with individual teachers at the beginning of the year, where the teachers had an opportunity to describe their own goals for the upcoming school year. At this time the principal would provide feedback and reinforcement to the teacher and also ensure that the teacher's goals were in some way linked to the mission of the school. One principal emphasized the importance of individual teacher growth:

> *So, in meeting with the teachers, I ask them to come up with at least one personal goal that would look at growth in regard to student centeredness. Therefore, over the course of the year, they are going to be making their classrooms a little bit better in terms of the instructional involvement that they have with students. I, as a school principal, may also request one school goal of them, which I shared with them in terms of growth. It has to do with growth.*

In describing their actions, principals stated strongly that instructional supervision was not something that just happened. It required initial thought, and then planning, organizing, and commitment in its undertaking. One principal in particular talked about organizing and planning on a school-wide basis.

> *Organization, planning, just the organization of the building, has to do with instructional supervision. All of those organization for instruction type things, timetables, have to do with instructional supervision.*

Although they tried to set expectations at the beginning of the school year, the principals recognized that this was insufficient in ensuring that they were followed throughout the year. Instead they actively monitored situations and where necessary intervened both positively and negatively to reinforce expectations. The positive reinforcement might be in the form of a note, or they might drop into a class and give a teacher words of praise for fine work. On the other hand, direct confrontation might have to be used to let the teacher know that "I cannot accept this." For example, one principal described a situation where a teacher yelled at a student and told the student to "shut up!" In reviewing the incident with the teacher, the principal made it very clear that treating students in that manner was not consistent with the established school goals. Another principal described a situation where he had to approach a teacher directly about her teaching methods. In this case, the teacher was using worksheets exclusively and this was not consistent with the school's approach to the teaching of language.

For these principals, instructional supervision was grounded in the purpose of the school, which had been agreed to by the staff at the beginning of the school year, but it was also constantly monitored and sustained through formal structures such as timetables and programs and informal discussions with teachers about what was happening in the school and what was appropriate.

Providing Resources and Support

Principals referred to a variety of ways in which they attempted to facilitate the work of teachers and the instructional process. As one principal stated, "I can't deliver curriculum but I do know that I can facilitate people who have ideas." In addition to the provision of materials and assistance, principals were willing to offer resources of a different kind. Support for the teachers was a primary way principals felt that they could assist teachers, and it included listening, providing extra time, and showing appreciation to teachers. One principal described a type of psychological and resource support that would reassure teachers that they were not alone when they were involved in some sort of change:

> *What you must provide them with is a tremendous backup of resource, advice, assistance, consultation, and so part of the model here of making change is to see that no staff member would be left with the feeling that they are by themselves in terms of understanding either their involvement with students or how to make more supportive involvement and change with their students. I operate on the basis that whatever advice, assistance, support, materials you need, you will get.*

One of the most immediate ways this could be done was by ensuring that teachers had the required materials available:

> *I think you have to make sure that the teachers have the tools to do the best job that they can and again, that'll vary with the school you're at and the people that are in the building as to how much of a priority you make that.*

Two of the principals referred to the act of allocating money as being a component of instructional supervision. One explained that budget allocations reflected whether the principal was more interested in capital acquisition than education.

Principals were also very willing to share their instructional expertise with the teachers but recognizing that they were not and could not be experts in all fields, they made arrangements to involve specialists, other teachers, or consultants. As one interviewee explained,

> *I say to the teacher, "Allow me to understand the student's concern and what you have got planned for him." I don't come to the teacher as the expert. I believe that the teacher knows more about the student's specific needs and concerns than I do.... I can offer some suggestions that I might have, but then I say, "I don't want you to worry about this. Help is on the way." The district has an abundance of support in consultants, specialists, materials.*

Time was also recognized as being a valuable resource for teachers. Principals described covering classes for teachers so that they could visit other schools or attend inservices. They would also bring in substitute teachers to enable their teachers to work on program development:

> *You bring in a half-day sub for the elementary teacher who's agreed to take on the junior high kids and develop a program. Give her half a day to do some curriculum development and clean up in the art room, get it organized, get ready, consolidate.*

Listening to teachers was described as being an important form of support. One principal referred to it as temperature taking:

> *Outside of that, to take the input that they give and listen to them, in the same way that you listen to kids. This just came back to mind as one of my philosophical approaches—the temperature taker. As an instructional leader, you're taking temperatures. There are all sorts of things you're taking the temperature of. Sometimes you ask people to take that temperature and feed the conditions back to you. If they run into some hot spots, well then they're to let you know about those.*

Principals also wanted teachers to know that they were valued and that their efforts were appreciated. This support might take the form of a personal note to the teacher, or as one principal described,

> *Sometimes it's not so much a direct comment as it is an action that you take that lets them know that they're valued. It might be a comment in assembly, it might be a direct one-on-one, it might be a suggestion that lets them know that they're pretty important.*

Two principals talked about creating a more conducive environment for instruction. In the first instance, the principal was referring to the physical environment of his school and described modifying open-area classrooms to make them closed classrooms and soundproofing rooms. In the second instance, the principal was referring to the school culture and described his task

as ensuring an appropriate standard of discipline to allow his teachers to teach:

> *Teachers in our schools are there to teach and if I can as principal create for them a climate which makes them more comfortable in meeting their daily challenges, then they will succeed instructionally. A most important aspect of instructional supervision is ensuring that the school climate is healthy. Teachers cannot be expected to be "policemen" and teach effectively at the same time. The principal, if need be, must assume the policeman's role in order that good instruction may be effected.*

Besides the role of disciplinarian, this principal also saw himself as a facilitator:

> *Through open and constructive communication these things that concern teachers and affect climate—poor student behavior, nonproductive use of time—become a responsibility of the principal. He is responsible for removing obstacles from teacher paths and helps teachers enjoy themselves and their work in the classroom.*

These principals thought that providing resources, materials, money, and time as well as access to consultants to their teachers was important. Equally important were everyday acts of listening, encouraging, suggesting alternatives, and of sustaining the school culture.

Observing Others and Being Observed

All the principals in this study mentioned that formal classroom observation was of primary importance in their role of instructional supervisor. However, the purpose of their involvement and the amount of time allocated to this activity varied with each principal. Many of the principals described the numerous demands on their time and varied widely in the time given to formal classroom observation. Those who placed priority on being in the classroom to see "what's happening," made a point of planning their visitations and writing them down. As one principal noted, "There's a lot of demand on principals' time. If you're going to do instructional supervision you've got to write it on your calendar." Although principals varied widely in the time given to formal classroom observation, all were involved in informal drop-ins as part of their daily walks around their schools.

To one principal, instructional supervision primarily meant "getting a sense of, feel for, instruction ... that the teachers are involved with the students." Accordingly, he ensured that "I am in everyone's classroom at least once a week. So that over the course of the year, I will be able to have a minimum of at least 40 to 50 visits with each teacher."

The school district required principals to write a number of formal teacher evaluation reports every year, and some of the principals made specific classroom visits for evaluation. For example, this is how one principal differentiated between a performance review and an evaluation. "Performance review, for myself, is a process of growth which means that everyone, regardless of whether they're a first-year teacher or have been in the field for 30 years, that each year is a year of growth."

However, if during a review visit the principal saw something he was concerned about, then, he considered it an evaluation which included the question of competency and the aim was to bring about necessary improvement:

The evaluation can take place at any time. The first day of classes or the last day of classes. But if a teacher interacts with a student in a way that is not supportive of our goal of student centeredness, then I will evaluate on the specifics of that involvement. That is separated from his performance review in terms of growth. If I question the academic program or the pro-social skill program that the teacher is involved in, with the student, then I will take that up specifically.

Another principal who similarly did not have separate evaluative classroom visits also arranged to observe teachers on a regular basis, but qualified that he would "always have some kind of focus determined" before the classroom visitation.

The focus would usually come out of an inservice activity that we were undertaking as a group. We may have been working on enhancing questioning skills or lesson design or some kind of a teaching strategy to do with peer teaching or something like that. And then with almost a contract between myself and the teacher, I would go in and give them feedback related to the skills or the strategies that we'd been talking about in the inservice, for example. But there were two sides to that contract. The other side was that the teacher would design a lesson in which they could demonstrate those skills.

Sometimes he would visit some teachers more than others and

Interestingly enough that was often because of a request from a teacher. And again usually related to something going on at an inservice. They knew that their next scheduled observation wasn't going to be until a month from now. They were working on something and they felt good about it and they wanted to share it so they invited me in to observe and we would do that.

One principal thought that one of the roles of an instructional supervisor was to be an effective model and strongly emphasized the importance of being a demonstration teacher. Being seen teaching served two purposes. It developed credibility with the staff and also enhanced the relationship with teachers by showing that the principal was "not perfect":

I don't think a principal can exercise instructional leadership unless they can demonstrate that they are also excellent teachers. So while I didn't have any formal regular teaching assignments, I did a lot of teaching. Very often with the teacher observing and later we could talk about what kinds of lessons were taught, teaching strategies, teaching skills—those kinds of things. So that's one role that I think is sometimes overlooked when you talk about instructional leadership and supervision in the school and that is the role of model.

Another principal, while acknowledging that teaching a class would make him a better curricular leader, felt that with all of the interruptions that were inherent in his job as principal, it would not be beneficial for the students to have him as their regular teacher. Unlike his colleague he did not seek out opportunities to provide demonstration lessons.

Other principals talked about less formal ways of visiting classrooms. In these informal visitations, the principals, in a sense, were not only observing, but were also being observed. They talked about "being visible" or "management by walking about." They explained that it let teachers and students know that they were interested and it also gave them a sense of what was happening in their school. They sought specific information when they

entered classrooms and some considered that principals should also be able to model teaching in order to encourage quality instruction.

One principal described how he took over the classroom duties of his teachers to allow them to observe another teacher. He saw this as being beneficial in two ways:

> *It conveys a belief in the professional competence of your staff. You're saying that you know that the teacher has the ability to give some feedback. I'm not the only one. It also demonstrates to the teacher that you know enough about real life circumstances because you're taking their class and that you're actually working with them.*

Another administrator relieved teachers for personal matters and viewed it as an opportunity to gain a sense of the classroom without the teacher being present:

> *You walk around and see what units teachers are working on, whether that's math, whether that's social studies.... and you will see what's on the walls, you'll see their plans, you'll see what is actually going on in the classroom so that you have a better knowledge of what really each teacher has.*

An informal approach used by a principal of a large high school, was to "be in classrooms a lot," but not with the purpose of evaluation. He did this to provide public support for the importance of classroom instruction to teachers and students. He depended on his department heads to "stay close to the day-to-day operation" and created a faculty council that met weekly to assist him in keeping in touch with what was happening throughout the school.

While all these principals thought it was important to be observed in classrooms because such visits stressed for students and staff the importance of instruction, only three made regular extended visits to classrooms in order to observe teachers and provide feedback on their teaching. Instead, most "popped in" to classrooms for brief periods on a fairly regular basis. Nor were they unaware of the difficulties in deciding on what actions should be taken when they had concerns about the instruction they observed. Although they had well developed district policies that they followed in the event of a major problem with classroom instruction, they were more hesitant in situations where the situation was less clear cut. As their conversations indicated, principals were constantly balancing their concern for the educational experiences of specific students and the concomitant necessity to insist that the teacher improve his or her instructional practices, with their desire to introduce specific changes to the work of the school and their need to have widespread staff support for the initiative. They were acutely aware that if they identified a teacher as "requiring improvement" they might well lose the positive regard of the remainder of the staff.

Encouraging Teacher Growth

One way principals chose to involve teachers was to develop strong professional development programs. One principal stated, "I feel that professional development is very important. That's where a large part of our budget is set aside." He described the importance of teacher involvement in a district program with an emphasis on enhancing teaching effectiveness. He recognized that it was expensive but felt that the benefits of discussing aspects of the instructional process with other colleagues were well worth it and noted that

everyone on his staff at present was involved with the program in some form. All the other respondents indicated some degree of involvement with this program. Many of them used it as a basis for their observation of teachers or to help clarify what teachers should be doing in the classroom. When describing his expectations of teachers, one interviewee said, "We're really teacher effectiveness people. We believe in it and we're always involved."

Another form of professional development used in one school was an extensive weekly program based on the model advocated by Joyce and Showers (1980). It involved an extended working day for the teachers four days per week to allow for early dismissal of the students on the day chosen for a professional development activity. The principal described the cycle of teacher involvement as: understanding a skill, demonstration of this skill, practicing the skill in a safe setting, and taking the skill into the classroom and practicing. His primary role in the process was providing coaching and feedback to the teacher practicing the skill in the classroom. He also attempted to free other teachers to perform the coaching role as well.

In addition to these school-based approaches to professional development, principals encouraged teachers to seek personal and professional growth. As one principal pointed out,

> *Part of my responsibility as an instructional supervisor is to see that the staff are stretched. Staff have to understand that their capacities are not finite. I do believe that it is my job to help them discern and rediscover talents, strategies, methods, and materials at these times in their endeavors when they feel that they have done their best and can do no better.*

Approaches mentioned were encouraging teachers to attend inservices, keeping the staff informed about professional development opportunities, having consultants come out to the school, freeing teachers to visit other teachers in their classrooms, or giving teachers on staff opportunities to share ideas with other staff members.

Encouraging the Involvement of Others

Principals actively pursued the involvement of others in the process of instructional supervision. To share some of their workload and at the same time involve and encourage knowledgeable people on staff, principals would delegate responsibilities. Two of the principals mentioned involving their assistant principals in the supervisory process. One principal mentioned that the involvement of his assistant principals allowed him the additional advantage of being able to work on "creative curriculum":

> *Having two assistant principals is a tremendous advantage to me. They each have their own particular areas.... So, from my point of view, it provides me with a tremendous channelling of responsibility and frees me to get involved in creative curriculum.*

Another principal delegated responsibilities to staff members to harness their skills and knowledge and to stimulate their professional growth:

> *Well, one thing is to take a look at each of the areas and delegate those responsibilities to other people. For instance, in the science area, to look at Steve and hold him accountable, in a sense, to review programs and to add to programs to make science a focus for this year for our budget process. So that's a leadership step. In other areas like health and*

phys. ed., to give Donna a task to find out what the resources are that are available within the school, relative to the list that's available to us—to see where the needs are. Doing that stimulates everyone to look at the curriculum and I guess what I'm saying there, again it's part of philosophy, is that you can't do it all yourself. It's impossible.

He emphasized that it was impossible for a principal to be knowledgeable in all of the curricular fields and while "You're certainly responsible for it ultimately," he considered it essential to build on the skills and expertise of others.

The involvement of others was not exclusive to teachers. Parents and the greater community and support staff were seen as being an important part of the instructional process as well:

As part of our statement of beliefs, you will recall that it says that the responsibility for teaching children lies with parents, teachers, and significant members in the community as well. And when we take a look at the instructional part, let's take a look at language arts. I don't expect the owner of the local Mac's store to be assisting Johnny with his reading skills, but I do expect the parents to be involved in that. When it comes to the area of pro-social skills, appropriate use of language, I believe that the teachers must assist Johnny in the appropriate use of language during school, the parents must teach Johnny the appropriate use of language in the home, and Mr. Owner of the Mac's Store must also assist with the appropriate use of language from Johnny when Johnny is in his store or in the community.

Another principal described involving parents in his school by having them assist teachers in the classroom and by making them a part of the school budget process.

Principals involved assistant administrators in formal evaluation procedures, and teachers in program responsibilities in their areas of expertise. Some involved parents and community members in helping sustain the work of the school.

Reflections

The four preceding categories presented the actions described by principals when they were asked the general question of how they acted as instructional supervisors. For these principals, instructional supervision encompassed much more than direct observation of teaching episodes. The work of the seven principals in the studies of Dwyer et al. (1985) was codified into nine categories: goal setting and planning; monitoring; evaluating; communicating; scheduling, allocating resources, and organizing; staffing; modeling; governing; and filling in for an absent teacher. All these activities were mentioned in this study. But the seven principals in Dwyer's case studies, like those in this study, gave meaning to these activities in differing ways through the articulation of "an overarching perspective of their school settings" (Barnett, 1986, p. 176). This perspective is usually referred to as culture, the values, beliefs, norms, rituals, ceremonials, and symbols that spell out the ways things are accomplished in an organization (Firestone & Wilson, 1984).

This perspective is also evident in this study. Underlying the categories is the notion of developing and sustaining a culture whose expectations would be the heartbeat of the school, one that would be sustained through talk and the everyday acts of the participants, hence the importance of the opening

activities when both a climate and specific philosophy were developed. The translation of the philosophy into a set of standards provided another means by which school members could judge their activities and was a ready reference point during discussions with them. Instructional supervision, therefore, was embedded in statements about the expectations for the teachers' orientation to children and curriculum. Having made such goals public, some principals also had teachers delineate their own expectations, shaping these through private conversations about the teacher's professional growth. Principals spoke about their expectations and how they were developed. The way they described their views reflected how they saw their purpose and role.

One principal stressed that staff should "adopt the mission" similar to the one he had outlined. Others talked about staff "adopting a statement of beliefs," or of needing teachers to "buy into the mission." Two principals talked about working with staff to "come up with" a vision, while one administrator emphasized the link between vision or mission and what happened in the school. He felt that he had "to make sure the staff realize.... our objectives are set from those." These quotes reveal that each of these principals had goals for education that were important to him. They also reveal how the principals went about involving their staffs in obtaining support for these goals and concurrently, the linkage of authority and communication in the culture of that school.

Another facet of culture was made manifest when these goals were monitored. Both by casual walkthroughs and detailed observations, the principal's presence was a visible reminder of the "contract," the stated goals or mission statement that teachers had supported. These findings confirm those of Dwyer (1984, 1986b) and Lee (1991) who found that principals attempted to shape their schools to fit a educative philosophy with which they were comfortable by shaping the meanings staff gave to their work.

Yet although these principals spoke about but did not name the development of a school culture, their use of words such as *buying* or *adopting* raise questions about the rights and responsibility of the principal in the process. For a number of these administrators it was sufficient that they design the mission statement that teachers would, by compliance or cajoling, adopt. Only two principals talked about spending time with their staffs to develop a statement of beliefs. Hence underlying the seemingly similar statements about strategies from the principals in this study, were profound differences in philosophy. While some principals stressed the instructional, educative function of the school and the importance of involving teachers as professionals who had expertise and expectations about how schools should be, other principals were more concerned with a school where there were clear expectations and benchmarks that could be used to monitor classroom instruction indirectly.

Important to some principals was the confirmation of their expertise through teaching, giving workshops, providing feedback on teaching skills, or providing advice about classroom situations. Such activities provided another forum through which principals sought to influence instruction. This area has received most emphasis in the recent literature especially in relation to ensuring that principals are competent in observing and assessing classroom in-

struction. The various reasons these principals had for popping into classrooms reflected the extent to which they saw themselves as experts and their greater concern over the need to sustain a cooperative, supportive school culture.

Greenfield (1986) analyzed the culture of one of Dwyer's case schools and identified three dimensions of the principal's work: social, moral, and technical. The social dimension encompassed the interpersonal component of the principal's work. In this study, all principals used talk as their major medium for shaping what happened in their schools, a confirmation of Dwyer's (1986b) conclusion that "Communication accounted for the largest proportion of each principal's actions" (p. 15), and of Wolcott (1973) and Gronn's (1983) earlier work. Blumberg (1989) raises questions about the focus of principals' talk, commenting that in a review of studies on the principalship, he had found that:

> Talk was the stock-in-trade. But what the talk was about was mostly left unstudied. The question then becomes, what is administrative talk about? This is another way of asking, What is the content of the work of administrators? (p. 43)

For these principals, the conversations frequently revolved around listening, problem solving, advice giving, the provision of resources and time, the linking of consultants and teachers, the reaffirmation of quality, and, ultimately, the sustaining of the culture.

The second dimension identified by Greenfield (1986) was moral, which he explained as the decision between varying conflicting standards of goodness inherent in every action whether deliberately reasoned or automatically chosen. The principals in this study were very conscious that actions carried meaning about what was right and good in the school setting. One example was the principal who admonished the teacher who was rude, stating that it was not in accordance with the goals of the school. Yet principals were also called on to make difficult moral decisions about whether to intervene in a classroom where a teacher who was close to retirement was providing poor quality instruction; whether to assign poor teachers to noninstructional support positions, thereby depriving competent teachers of the opportunity; whether to challenge a popular but lazy teacher; or whether to risk alienating the staff and jeopardizing an important program change by identifying a teacher as needing "improvement." Like the principals in Dwyer et al.'s (1985) case study, the principals in this study chose to ignore those system directives about extensive classroom observation where they felt it was unwarranted.

Greenfield (1986) subdivided his third dimension, technical demands, into organizational and instructional. While the organizational demands related to planning, decision making, and goal setting are clearly evident in this study, the principals varied in the amount of time they allocated to the central instructional dimensions of curriculum development and articulation, and improvement of instruction. Some of this can be related to the grade level and size of the school, some to the principals' own admission that they were not curriculum or instruction experts.

Greenfield (1986) concluded:

> While concepts like instructional leadership, management, and administration dominate discussions, descriptions, and interpretations of the work of school principals, they provide little leverage on understanding the meaning of or even adequately describing the nature of a principal's actual work activities and behaviors. They are ambiguous concepts, and.... they do not adequately describe what principals actually do, nor do they help one decipher the meaning of those actions in a given context. (p. 130)

Similarly, Daresh (1991) revealed that, for him, principals acted holistically rather than in response to a set of tasks or procedures. The principals in this study did not view their supervisory activities as separate from their everyday activities as principals. In fact, it is this embeddedness that militates against instructional supervision having a discrete meaning for these principals.

References

Barnett, B.C. (1986). Principals creating case studies of one another: The peer-assisted leadership program. *Peabody Journal of Education, 63,* 174-186.

Blumberg, A. (1989). *School administration as craft. Foundations of practice.* Boston, MA: Allyn and Bacon.

Costa, A., & Garmston, R. (1985). Supervision for intelligent teaching. *Educational Leadership, 42,* 70-81.

Daresh, J. (1991). Instructional leadership as a proactive administration process. *Theory into Practice, 30,* 109-118.

Dwyer, D.C. (1984). The search for instructional leadership: Routines and subtleties in the principal's role. Educational Leadership, 41, 32-37.

Dwyer, D.C. (1986a). Frances Hedges: A case study of instructional leadership. *Peabody Journal of Education, 63,* 19-86.

Dwyer, D.C. (1986b). Understanding the principal's contribution to instruction. *Peabody Journal of Education, 63,* 3-18.

Dwyer, D.C., Lee, G.V., Barnett, B.G., Filby, N.N., Rowan, B., & Kojimoto, C. (1985). *Understanding the principal's contribution to instruction: Seven principals, seven stories. Vols. 1-8.* San Francisco, CA: Far West Laboratory.

Firestone, W.A., & Wilson, B.L. (1984). What can principals do? Culture of school is key to more effective instruction. *NASSP Bulletin, 68,* 7-11.

Glickman, C.D. (1985). *Supervision of instruction—A developmental approach.* Boston, MA: Allyn and Bacon.

Greenfield, W.D. (1986). Moral social and technical dimensions of the principalship. *Peabody Journal of Education. 63,* 130-149.

Gronn, P. (1983). Talk as the work: The accomplishment of school administration. *Administrative Science Quarterly, 28,* 1-21.

Joyce, B., & Showers, B. (1980). Improving inservice training: The messages of research. *Educational Leadership, 37,* 379-385.

Lee, G.V. (1991). Instructional supervision as collaborative sense-making. *Theory into Practice, 30,* 83-90.

McEvoy, B. (1987). Everyday acts: How principals influence development of their staffs. *Educational Leadership, 44,* 73-77.

Morris , V.C., Crowson, R.L., Porter-Gehrie, C., & Hurwitz, E. (1984). *Principals in action: The reality of managing schools.* Columbus, OH: Merrill.

Oliva, P.F. (1984). *Supervision for today's schools* (2nd ed.). New York: Macmillan.

Owens, R.G. (1991). *Organizational behavior in education* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Pajak, E. (1990). Dimensions of supervision. *Educational Leadership, 48,* 78-80.

Sergiovanni, T.J. (1987). *The principalship: A reflective practice perspective.* Newton, MA: Allyn and Bacon.

Smith, W.F., & Andrews, R. L. (1989). *Instructional leadership: How principals make a difference.* Alexandria, VI: ASCD.

Smyth. W.J. (1988). *A critical pedagogy of teacher evaluation.* Geelong, NSW: Deakin.

Wiles, J., & Bondi, J. (1986). *Supervision. A guide to practice* (2nd ed.). Columbus, OH: Merrill.

Wolcott, H.F. (1973). *The man in the principal's office: An ethnography.* New York: Holt, Rinehart & Winston.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 2, June 1992, 121-132*

Douglas J. Smith

and

Robert P. Sanche
University of Saskatchewan

Saskatchewan Interns' Concerns at Three Stages of a Four-Month Practicum

The present study was undertaken to document the concerns of preservice teachers during a four-month practicum and to determine if the nature and intensity of concerns were consistent with the developmental sequence postulated by Fuller and Bown (1975). One hundred and twenty-five interns placed in schools during the winter of 1990 were the subjects for the study. Mean concerns scores for survival, task *and* pupil *clusters on the Intern Concerns Checklist decreased significantly during the internship. In keeping with Fuller's findings, survival concerns decreased most over time. In contrast to Fuller's findings, the present sample of interns had consistently greater concern for pupils and showed those concerns earlier in their internship than had subjects in earlier studies. The current topics and sequence for the internship seminars generally match the emerging concerns of interns, although there may be need for an earlier treatment of content related to* pupil *concerns.*

La présente étude avait pour but d'explorer les préoccupations des enseignants en formation pendant un stage pratique de quatre mois, et de déterminer si la nature et l'intensité des préoccupations étaient en accord avec les phases de développement postulées par Fuller et Bown (1975). Cent vingt-cinq stagiaires placés dans les écoles pendant l'hiver 1990 ont servi de sujets à l'étude. Les moyennes des scores pour les thèmes survie, tâche, *et* élève *du Intern Concern Checklist ont baissé d'une manière significative pendant le stage. En conformité avec les résultats de Fuller, les préoccupations concernant la survie ont diminué le plus. Contrairement aux résultats de Fuller, le présent échantillon de stagiaires a montré plus de préoccupations pour les élèves, et cela plus tôt durant leur stage, comparativement aux résultats des études antérieures. Les contenus actuels et l'ordre des séminaires de stage étaient, d'une manière générale, en accord avec les préoccupations des stagiaires. Cependant, il ya peut-être le besoin d'un traitement plus prématuré du contenu concernant les préoccupations au sujet des* élèves.

Faculty in the College of Education, University of Saskatchewan, are in the third year of intensive discussion and revision of their teacher education programs. As a central and unifying component of the programs, the four-month "extended practicum" or internship, which comes in the latter part of the programs, has been updated annually. Ideas from the published literature on effective teacher education practices and theoretical concepts about what

Douglas Smith is an assistant professor in the Department of Curriculum Studies. His interests include teacher supervision, field experience, and models of teaching.
Robert Sanche is a professor in the Department for the Education of Exceptional Children. His interests include special education service delivery, mentoring relationships, and field practicum supervision.

constitutes a reflective practitioner (Richards & Gipe, 1987; Schon, 1987) are factors that are also being reexamined. In the review currently in progress, both theoretical and research-based views of how interns learn and develop during the four-month internship are being considered. Our own students are providing data to test existing theories of intern growth and to guide program changes.

One approach to monitoring growth during practica has been that of measuring changes in the concerns expressed by interns during the process (Fuller, 1969; Fuller & Bown, 1975). The term *concerns* can be characterized as the "anxiety or anxious sense of responsibility" that interns or beginning teachers feel as they are initiated into the profession. According to Fuller's formulation, although optimum levels of concern provide a natural impetus toward learning, excessive levels of concern are detrimental to intern learning and progress.

In her earlier work, Fuller (1969) postulated that beginning teachers progress through an initial preteaching stage in which they identify with their pupils and are not concerned about the specific demands of teaching. Teachers then proceed through an early stage of concern about *survival* and the specific elements of the teaching position, to a later stage of concern for the learning of their pupils. In her more recent conceptualization (Fuller & Bown, 1975) Fuller refined her original theoretical formulation to a three-stage progression of concerns. Stage 1 was characterized as the time of transition from student to teacher, when the cluster of student concerns relates to personal adequacy and survival as a teacher. The novice teacher then progresses to stage 2, in which the focus of concern is a set of teaching *tasks* that have to be mastered. At Stage 3, concerns of the new teacher are related to tailoring content to individual pupil needs, to responding to the social and emotional needs of pupils, and to being fair. In this developmental model, it is essential that the new teacher overcome survival concerns before it is possible to adequately address task and pupil concerns (Fuller & Bown, 1975). Griffin et al. (1983) noted that, besides these changes in locus of student concerns from survival to pupil, there are also changes in the intensity of student concerns within each of these clusters.

There is some evidence in recent research to support Fuller's (Fuller & Bown, 1975) conceptualization of a developmental sequence of teachers' concerns as a valid framework for monitoring student growth during internships. Holmes and Barbour (1988) conducted structured interviews with students after reviewing their ongoing progress records and journal entries. The authors concluded that during teacher preparation, greater attention needs to be given to the developing priorities and concerns of beginning teachers. Richards and Gipe (1987) reviewed reflective statements in interns' journals and reported evidence supporting the concept of a developmental sequence of intern concerns. O'Sullivan and Zielinski (1988), in studying the validity and reliability of a "stages of concern" checklist used with 92 fifth-year preservice teachers, found that the progression of expressed concerns approximated that of Fuller's developmental stages. Butler and Smith's (1989) interns had significantly different developmental concern levels as they progressed through their year-long internship, and Hall and Jones (1976), Loucks (1977), and Ryan

(1979) all found evidence of a sequence of developmental concerns for their subjects. Berliner (1988) and Bercik and Blair (1989) recommended that both intern education and evaluation should be provided in a manner consistent with the relevant stage of intern development. Studies by Sitter and Lanier (1982) and Zeichner (1984), however, did not provide support for the Fuller and Bown (1975) concept of a developmental sequence of intern concerns.

At the University of Saskatchewan the protocol for the extended practicum is predicated on an implied developmental model of intern competencies. Nevertheless, neither the stages of development nor the concerns of interns have ever been documented. Further, no research has been carried out to determine whether the sequence and content of the internship experiences correspond with intern needs and concerns during the process.

The objectives of the present study are to determine:

1. the actual concerns of Saskatchewan interns at three times during the extended practicum,
2. if the concerns of individuals and groups of interns change in a manner similar to that postulated by Fuller and Bown (1975), and
3. if the current sequence of internship seminar topics relates to the evolving concerns of interns.

The Saskatchewan Internship Program

The extended practicum at the University of Saskatchewan was developed over a decade ago and has been updated and refined several times over the years. At the present time it consists of two elements: a progression of effective teaching competencies expected to be learned by all students and a sequence of clinical procedures designed to provide daily supervision by school-based cooperating teachers. Monthly supervision is provided by College of Education faculty and staff. A detailed Internship Manual includes the philosophy, sequence of skills, and an overview of inservice and supervision activities for interns. Interns are also required to maintain daily Planning and Observation books, which are routinely reviewed by the College of Education supervisors.

A full-day internship seminar is held for all intern-cooperating teacher pairs during the first week of the extended practicum. Three subsequent seminars, conducted by college supervisors, are held at approximately one month intervals during the internship. The teaching competencies highlighted during the seminars are those considered of contemporary relevance for the interns' developmental stages. Between internship seminars, cooperating teachers provide daily clinical supervision of the competencies presented in the most recent seminar. The college supervisors provide between seminar rounds of clinical supervision for each intern. As well, they rely on the cooperating teachers' clinical supervision reports in the Planning and Observation book to judge the progress of each intern. An overriding objective of clinical supervision is to help the intern to reflect continually about his or her teaching.

The content of the internship seminars is sequenced to approximate the anticipated developmental needs and concerns of interns. During the first seminar the emphases are on: (a) the identification of professional attributes, (b) the establishment of a productive working relationship between intern and cooperating teacher, (c) individual and joint expectations of intern-teacher

pairs, (d) basic classroom management skills for interns and (e) lesson and unit planning and delivery. The second seminar is focused on extending the interns' repertoire of classroom management skills and on keeping accurate records of teaching progress during clinical supervision. In the third seminar, questioning and responding behaviors are featured, and the final seminar is focused on lesson presenting skills such as teacher clarity and enthusiasm. Issues related to the interns' two weeks of full-time teaching and to the final internship evaluation are also discussed at the last seminar. As has been mentioned earlier, since the program began no research data have been gathered to document either the implied developmental sequence of intern growth or the concerns interns have as they progress through the four-month internship.

Methods

Sample

The sample for the present study consisted of 152 elementary and secondary education interns placed by the College of Education for the spring semester, 1990. One hundred and three of the interns were females and 49 were males. Sixty-six were placed in rural schools and 86 in urban schools. Interns who subsequently dropped out of the extended practicum or who failed to complete one of the three administrations of the Teacher Concerns Checklist were dropped from the study. At the end of the extended practicum, complete data were available for 127 (41 males and 86 females) of the original sample.

Instrumentation

The 50-item Teacher Concerns Checklist developed by Fuller and Borich (1988) was used in the present study. It measures teacher survival, task, and pupil concerns and employs a Likert-type five-point scale ranging from (1) not concerned to (5) totally preoccupied. The survey prompts interns to recall and rate current levels of concern at each administration.

The Teacher Concerns Checklist (TCC) items were constructed from the actual responses to 1,359 open-ended teacher concerns statements and from items derived from factor analysis of the open-ended concerns' statements (Fuller, Parsons, & Watkins, 1974). The TCC has reliability coefficients of internal consistency ranging from .79 to .91 (N=335) and one-week, test-retest stability coefficients ranging from .77 to.87 (N-44) (Parsons & Fuller, 1974).

Through Varimax rotated factor analysis of the open-ended concerns of two additional teacher samples (N=1,028 preservice and 265 inservice; N=335 preservice and 345 inservice teachers), three factors, survival, task, and impact-on-pupil, received greatest weight. The three factors, survival, task, and pupil concerns, did load when McBride (1985) had the data from 30 physical education inservice teachers subjected to a maximum likelihood factor analysis. In order to establish the suitability of the Teacher Concerns Checklist for use with interns at the University of Saskatchewan, we initiated validation procedures with College of Education internship supervisors. All these staff had supervised interns four or more terms. Ten of these supervisors independently categorized the 50 checklist items as measuring either survival, task, or pupil concerns. Only items categorized by majority agreement of the 10 college supervisors were included. Ten items that either could not be

categorized or were categorized differently by a majority of the 10 "experts" were dropped and the 40 remaining items formed the survival, task and pupil concern clusters. Thirty-two of the 40 remaining items had also been validated through factor analysis at Western Montana College (Rogan, 1988).

Procedures

The revised (40 items) Teacher Concerns Checklist (hereafter called the Intern Concerns Checklist) was completed by interns during the first, eighth, and 16th weeks of the extended practicum. At the end of the first internship seminar the interns completed the Intern Concerns Checklist and provided demographic data on age, gender, program, college year, rural/urban placement for internship, rural/urban school from which each graduated, grade level of placement, area of specialization and average number of students in the assigned classroom.

Data Analysis

Mean subscale scores were calculated for all respondents for each administration of the Intern Concerns Checklist. Changes in mean scores over the four-month internship were assessed using a separate one-way ANOVA with repeated measures for each of survival, task, and pupil subscales (see Table 1). The SPSS program MANOVA was used to determine changes in subscale mean scores over time. It was also used to test the sphericity of the data and the significance of the differences between subscale means for each set of subscale scores.

Frequencies, percentages, and modes were calculated for survival, task, and pupil subscales of the Intern Concerns Checklist for each of the three administrations. Model scores for subscale clusters of the checklist were plotted to show changes in intern concerns (see Tables 2, 3, and 4). In order to assess the level of concern of interns on individual checklist items at the beginning, middle and end of the practicum, the researchers arbitrarily chose a model score of 4 as the minimum score to indicate significant concerns (see Tables 2, 3, and 4).

Results and Discussion

Intern Concerns Checklist Findings

Interns' mean ratings for survival, task, and pupil concern subscales at the first, second and third checklist administrations are presented in Table 1. Mean concerns ratings for each of the three subscales decreased significantly ($p>.001$)

Table 1
Mean Intern Concerns Subscale Scores by Time of Administration

	Time of Administration					
	1		*2*		*3*	
Concern Subscales	*Means*	*SD*	*Means*	*SD*	*Means*	*SD*
Survival	3.17	.61	2.67	.63	2.46	.68
Task	2.44	.50	2.26	.52	2.23	.61
Pupil	3.15	.59	2.98	.59	2.87	.69

N = 128

Table 2
Individual Concerns and Group Modal Concerns for Survival by Subscale Item

Survival Subscale Items	Time of Administration 1 Checklist Concerns[a] Modes	%	2 Checklist Concerns Modes	%	3 Checklist Concerns Modes	%
Getting a favorable evaluation	4	60[b]	4	42	3	34
Doing well when supervisor present	4	52	3	33	2	37
Accepted and respected by my peers	3	51	3	38	2	35
Maintain class control	4	50	3	48	2	36
Students who disrupt the class	4	48	3	41	3	38
Feeling adequate as a teacher	4	48	2	35	2	36
Being liked by students	2	42	1+2	37	2	39
Feeling under pressure	3	40	2	40	2	33
Clarifying limits of authority and responsibility	3	40	2	40	2	33

[a]N=127 respondents on the Concerns Checklist
[b]Percentage of respondents who chose the mode

during the 16-week internship. The specific results for each of the subscales were: survival $F(2,254)=93.54$, $p<.001$; task $F(2,254) = 12.78$, $p<.001$; pupil $F(2,254) = 15.04$, $p<.001$. O'Sullivan and Zielinski (1988) also found an overall decline in intensity of concerns between early and later measures of concerns among their undergraduate students.

Survival

Mean survival concerns decreased over time (see Table 1) more extensively than did task or pupil concerns. The modes and percentages of interns having survival concerns also decreased over time (see Table 2). The decrease in numbers of interns having concerns early and late in the internship can be seen in the percentage columns of Table 2. Initially 60% of the total number of interns had high concern for the survival item "getting a favorable evaluation," whereas at the final checklist administration only 34% of the interns had the same levels of concerns. These results are consistent with Adams' (1982) and Maxie's (1989) research and Fuller's theory that interns are initially concerned about survival factors but become less concerned about these factors over time. According to Maslow (1968), people (interns) will initially feel threatened and inadequate in their new situations before becoming comfortable enough to move on to other concerns. Maxie (1989) also speculated that the positive supervision practices of the cooperating teachers could be associated with the decrease in survival concerns of interns.

Being supervised, however, was not always viewed with favor by Saskatchewan interns. "Getting favorable evaluations of their teaching" was a highly rated survival concern both initially and at the mid-term of the internship. This concern may be an expression of changes in role status as interns moved from student to teacher.

Saskatchewan interns had initial high survival checklist item concerns about maintaining class control (see Table 2). This finding parallels that of Campbell and Green (1990) who found that their interns were initially concerned about classroom control and possible chaos during teaching. Fuller and Bown (1975) found that initial concerns about class control were most intense for their preservice teachers. Veenman (1984), reviewing 83 international studies, also identified classroom discipline as the most frequent problem of beginning teachers.

Task

Fuller and Bown's (1975) theory that concern for task is the second stage in the evolving concerns of interns is not supported by the mean concerns scores (see Table 1). There was little change in concerns for task between initial, midpoint, and final checklist administrations; however, interns were more intensely concerned with the task items: "increasing my proficiency in content," and "selecting and teaching content well" (see Table 3). These concerns may indicate a need for coaching interns to enable them to transfer from preservice theory and content to the specific realities of the daily classroom (Joyce & Showers, 1981).

Pupil

The significant concern for pupils, measured at the first checklist administration was unexpected (see Tables 1 and 4). Fuller and Bown (1975) did note that preservice teachers express deep concern about pupils' needs and about relating to pupils as individuals, but that preservice teachers' survival con-

Table 3
Individual Concerns and Group Modal Concerns for Task by Subscale Item

	Time of Administration					
	1		*2*		*3*	
Task Subscale Items	*Checklist Concerns*[a]		*Checklist Concerns*		*Checklist Concerns*	
	Modes	*%*	*Modes*	*%*	*Modes*	*%*
Too many noninstructional duties	1	65[b]	1	60	1	48
Presenting all required material	4	63	3	42	3	42
Ineffective staff meetings	1	62	1	55	1	50
Too many students in class	1	58	1	55	1	42
Inadequate assistance from specialized teachers	1	57	1	63	1	54
Working with too many students each day	1	54	1	51	1	45
Increasing my proficiency in content	4	51	4	43	3	34
Nature and quality of instructional materials	3	46	3	33	3	36
Too many standards and regulations set	2	46	2	40	2	42
Rapid curriculum and instruction change	2	44	2	36	1	41
Lack of freedom to initiate	2	42	1	46	1	50
Selecting and teaching content well	4	41	4	52	4	38
Lack of instructional materials	3	36	2	37	1	38

[a]N=127 respondents on the Concerns Checklist
[b]Percentage of respondents who chose the mode

Table 4
Individual Concerns and Group Modal Concerns for Pupil by Subscale Item

Pupil Subscale Items	Time of Administration 1 Checklist Concerns[a] Modes	%	2 Checklist Concerns Modes	%	3 Checklist Concerns Modes	%
Ensuring students grasp content	4	65[b]	4	41	4	42
Increasing students' feeling of accomplishment	4	63	4	54	4	59
Meeting needs of different students	4	59	2	40	4	43
Whether each student is getting what he needs	4	54	4	42	3	42
Challenging unmotivated students	4	53	3	46	3	38
Whether students can apply what they learn	4	52	3	42	3	50
Guiding students toward intellectual and emotional growth	4	52	4	38	3+4	34
Slow progress of certain students	3	52	3	46	3	42
Adapting to needs of different students	3	50	3	38	3	40
Helping students to value learning	4	50	3	41	3	40
Recognizing social and emotional needs of students	4	49	3	40	3	50
Wide range of student achievement	3	46	2	40	3	40
Diagnosing student learning problems	3	43	3	39	3	38
Being fair and impartial	3	41	4	34	4	31
Diversity of student ethnic and socioeconomic backgrounds	3	39	2+3	31	1	34
Chronic absence and dropout	1	32	1	34	2	33
Student health and nutrition	3	28	2	40	2	33

[a]N=127 respondents on the Concerns Checklist
[b]Percentage of respondents who chose the mode

cerns overshadow these pupil concerns until such teachers have learned to cope with survival tasks. They postulated that expressed concern for pupils is manifest in a final developmental stage. The high levels of concern for pupils in the early stages of internship means that the Saskatchewan interns were perhaps less preoccupied with survival than had been hypothesized. The greater initial concern for pupils may also have reflected a more pupil-centered preservice education program at the University of Saskatchewan than that experienced by the Fuller and Bown (1975) sample. It must be noted, however, that concern for pupils was still high at the end of the internship (see Table 1). This continued high concern for pupils offers some support for the position of Fuller and Bown.

Saskatchewan interns coped with survival concerns while maintaining a high level of pupil concerns throughout their internships. Interns continually wanted to "ensure that students grasped subject matter content" and to "increase the students' sense of accomplishment" (see Table 4). Because this sample completed internship in the winter term, cooperating teachers would already have known each pupil and have had classroom discipline established. If this information about pupils was shared with interns, it may have

influenced their early concern for pupils. Other researchers attribute high measured pupil concern to supervising teachers' modeling and to the fact that interns want to do what supervisors expect of them (Adams & Martray, 1980).

Demographic Items and Intern Concerns

When interns' concerns were analyzed according to demographic factors, class size and grade placement level were significant. Application of Chi-square revealed that, near the end of the internship, interns with moderate class size (16-25 students) had significantly stronger survival and pupil concerns ($p>.05$) than did interns with smaller or larger class sizes. Larger-sized classes would normally be presumed to increase interns' survival concerns. In this sample, however, the small number of interns with larger classes made comparison problematic.

Interns in middle years (grades 7-9) placements had significantly higher task concerns ($p>.05$) than interns in elementary school (grades 1-6) placements. This finding may reflect a lack of specific middle-years task emphasis during preservice education. There were no significant differences for groups on other demographic variables.

Internship Seminar Topics and Interns' Concerns

The third purpose of the present study was to determine if the current sequence of internship seminar topics is consistent with the emerging needs and concerns of interns. First, a comparison was made between the actual content of the seminars and the developmental sequence of interns' concerns postulated by Fuller and Bown (1975). Then a comparison was made of the actual content of the seminars and the expressed concerns of interns at the beginning, middle and end of the extended practicum. In reviewing the earlier presentation of the content and sequence of seminar topics (see the section on "The Saskatchewan Internship Program"), it seems clear that the seminar emphasis changes from an initial focus on survival skills to that of task by midterm and finally to a focus on the needs of the pupil. Although there appears to be a good correspondence between the actual content and sequence of the seminars and the theoretical concerns of interns, there remains the question of whether their expressed concerns on the Intern Concern Checklist are consistent with the content and sequence of seminar topics (see Tables 2, 3, and 4).

During the first week of the extended practicum students completed the Intern Concerns Checklist, which yielded modal scores of 4 for 5 of the 9 survival items, thereby providing corroboration for the early emphasis on survival skills at the initial seminar. Interns also provided modal scores of 4 for 3 out of 13 task items and for 8 out of 17 pupil items. As stated earlier, these expressed intern concerns for pupils so early in the practicum were not anticipated. Certainly the content of the first two seminars was not related to pupil concerns. One possible reason why the Saskatchewan interns were developmentally more advanced in their initial concern for pupils may have been that all of them had completed either a two- or three-week full-time round of practice teaching in the year prior to their extended practica. Because all of them had that background of experience, perhaps more content in the initial two seminars could have been devoted to interns' concerns for pupils.

At the time of the third internship seminar, only one of the survival items, "getting a favorable evaluation," had a modal rating of 4, while two of the task and five of the pupil items also had ratings of 4. The emphasis on task and pupil concerns in the third seminar seemed moderately relevant for these interns, since task concerns had already receded in significance for them. By the time of the fourth seminar, no survival concerns, one task concern, and five pupil concerns rated modal scores of 4. Because the content of the fourth seminar was mainly on lesson presenting skills and pupil concerns, it seems clear that the content was appropriately placed for the group, even though many of the earlier pupil concerns had dissipated.

Intern Concern Checklist items that persisted at modal ratings of 4 to the final seminar may require further attention. For example, the persistently highly rated task item "selecting and teaching content well" may indicate the need for further seminar content about lesson and unit planning or new seminar content about models of teaching.

Although there is strong evidence that the content and sequence of internship seminars are consistent with the theory of Fuller and Bown's (1975) developmental model, the evidence is more tentative when we consider interns' expressed concerns on the Intern Concerns Checklist. Apparently Saskatchewan interns are developmentally advanced in their concerns for, and perhaps readiness for, an earlier treatment of task and pupil related teaching skills. Perhaps greater consideration should also be given to making the earlier, short-term practice teaching experience an integral part of the developmental progression of intern growth.

Conclusion

Concerns of interns at the University of Saskatchewan cluster around the survival, task, and pupil categories postulated by Fuller (1969). Like Fuller's subjects, these interns experienced a strong initial concern for survival, which diminished over time. In contrast to Fuller and Bown's (1975) developmental theory, Saskatchewan interns had strong concerns for pupils initially and throughout the practicum. Such early concern for pupils may have been partly a result of prior student teaching experiences and partly a result of other unknown elements of the preservice teacher education program at the University of Saskatchewan. These findings are consistent with those of Hord, Rutherford, Huling-Austin, and Hall (1987), who noted that at any given time some degree of survival, task, and pupil concerns exist in beginning teachers during all stages of development.

Although the content and sequence of internship seminars is generally consistent with the developmental concerns of Saskatchewan interns, there is need to consider content related to pupil concerns earlier in the four-month progression. The mandatory two- or three-week practice teaching experience prior to the extended practicum may allow for a better rationalization of the current progression of seminar topics and content.

There is need to design a study in which Saskatchewan interns articulate their unique concerns three or four times during the Extended Practicum. Because the Saskatchewan sample of interns appears to be developmentally different from those of earlier studies, using the Intern Concerns Checklist

may be a limitation in such a study. Any Saskatchewan intern's concerns not reflected in the current 40 items of the scale will go unreported.

Alternative methods of collecting interns' concerns should also be considered. Interviewing interns, or having them keep reflective journals, are potentially productive ways of identifying emerging concerns during the extended practicum.

Although the Intern Concerns Checklist has proven to be a useful tool for documenting the concerns of Saskatchewan interns and reviewing the content and sequence of practicum seminars, it seems clear that it has some limitations. There is need to devise more precise tools to monitor the developmental growth of interns during the extended practicum. Further refinement of practicum experiences will be enhanced by the development of such tools.

References

Adams, R. (1982). A look at changes in teacher perceptions across time. *Journal of Teacher Education, 33*(4), 40-43.

Adams, R., & Martray, C. (1980). *Correlates of teacher perceived problems*. Paper presented at the 9th annual conference of the Mid-South Educational Research Association, New Orleans. (ERIC Document Reproduction Service No. ED 195 567)

Bercik, J., & Blair, S. (1989). *A perspective of new teachers and their induction year at a major urban parochial school system*. (ERIC Document Reproduction Service No. ED 304 406)

Berliner, D. (1988). *The development of expertise in pedagogy*. Paper presented at the American Association of Colleges for Teacher Education, Washington, DC. (ERIC Document Reproduction Service No. ED 298 122)

Butler, E., & Smith, D. (1989). *A study of factors associated with fifth-year teacher interns' concerns, problems and stress*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Campbell, C., & Green M. (1990, June). *Becoming a teacher: The preservice component*. Paper presented at the Canadian Society for the Study of Education, Victoria, BC.

Fuller, F.F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal, 6*(2), 207-226.

Fuller, F.F., & Borich, G. (1988). Teacher concerns checklist. In G. Borich (Ed.), *Effective teaching methods*. Toronto, ON: Merrill.

Fuller, F.F., & Bown, O.H. (1975). Becoming a teacher. In K. Ryan (Ed.), *Teacher education* (Seventy-fourth Yearbook of the National Society for the Study of Education, pp. 25-52). Chicago, IL: University of Chicago Press.

Fuller, F.F., Parsons, J.S., & Watkins, J.E. (1974). *Concerns of teachers: Research and reconceptualization* (UTR&D no. 2323). Austin, TX: University of Texas Research and Development Center for Teacher Education. (ERIC Document Reproduction Service No. ED 091 439)

Griffin, G., Barnes, S., Hughes, R., O'Neal, S., Defino, M., Edwards, S., & Hukill, H. (1983). *Clinical preservice teacher education: Final report of a descriptive study* (Report No. 9026). Austin, TX: University of Texas, Research and Development Center for Teacher Education.

Hall, G.G., & Jones, H.L. (1976). *Competency-based education*. Englewood Cliffs, NJ: Prentice-Hall.

Holmes, E., & Barbour, C. (1988, February). *Transitions to teaching: A descriptive study of the priorities of teachers-in-training*. Paper presented at the annual conference of the Eastern Educational Research Association, Miami Beach, FL.

Hord, S., Rutherford, W., Huling-Austin, L., & Hall, G. (1987). *Taking charge of change*. Reston VA: Association of Supervision and Curriculum Development.

Joyce B., & Showers, B. (1981). Transfer of training: The contribution of coaching. *Journal of Education, 163*(2), 163-172.

Loucks, S. (1977). *Concerns expressed by elementary school teachers about the implementation of the SCIS curriculum*. Paper presented at the annual meeting of the Association for the Education of Teachers of Science, Cincinnati, OH. (ERIC Document Reproduction Service No. ED 250 163)

Maslow, A. (1968). *Toward a psychology of being*. New York: Van Nostrand Reinhold.

Maxie, A. (1989). *Student teachers' concerns and the student teaching experience: Does experience make a difference?* Paper presented at the annual meeting of the American Educational Research Association. (ERIC Document Reproduction Service No. ED 308 164)

McBride, R. (1985). *The concerns of inservice physical education teachers as compared with Fuller's concern model*. Paper presented at the National Convention of the American Alliance for Health, Physical Education, Recreation and Dance, Atlanta. (ERIC Document Reproduction Service No. ED 265 122)

O'Sullivan, K., & Zielinski, E. (1988). *Development of a stages of concern questionnaire for preservice teachers*. Paper presented at the Annual Meeting of the National Association for Research in Science teaching Lake of the Ozarks, MO. (ERIC Document Reproduction Service No. ED 295 811)

Parsons, J.S., & Fuller, F.F. (1974). *Recent research on two assessment instruments*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago. (ERIC Document Reproduction Service No. ED 093 987)

Richards, J., & Gipe, J. (1987). *Reflective concerns of prospective teachers in early field placement*. Paper presented at the Annual Meeting of the American Educational Research Association. (ERIC Document Reproduction Service No. ED 296 976)

Rogan, J. (1988). *Validation of stages of concern checklist*. Personal correspondence with author.

Ryan, K. (1979). *Some feedback is better than others: Implications of a study of first year teachers for the follow-up of teacher education graduates*. Paper presented at the meeting of the American Educational Research Association, San Francisco.

Schon, D. (1987). *Educating the reflective practitioner*. San Francisco, CA: Jossey-Bass.

Sitter, J.P., & Lanier, P.E. (1982, March). *Student teaching: A stage in the development of a teacher?* Paper presented at the annual meeting of the American Educational Research Association, New York.

Veenman, S. (1984). Perceived problems of beginning teachers. *Review of Educational Research, 54*(2), 179-223.

Zeichner, K.M. (1984, January). *The ecology of field experience: Toward an understanding of the role of field experience in teacher development*. Paper presented at the annual meeting of the Association of Teacher Educators, New Orleans.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 2, June 1992, 133-140*

Murray Sandell

and

Keith Sullivan
Dalhousie University

Teacher Disillusionment and Supervision as a Part of Professional Development

Approximately 70% of the 42 teachers in this study personally experienced disillusionment with their profession. The causes of disillusionment are analyzed in this study by three levels of professional development, namely, professional maturing, the consolidated learner, and the embryonic teacher. It was found that the causes of disillusionment and the advice to supervisors specific for assisting disillusioned teachers differed through the three levels of professional growth.

Environ 70% des quarante-deux enseignants dans cette étude ont personnellement exprimé un désillusionnement avec leur profession. Les causes du désillusionnement ont été analysées dans cette étude à partir de trois niveaux de développement professionnel, à savoir, la maturité professionnelle, l'apprenant consolidé, et l'enseignant embryonnaire. On a trouvé que les causes du désillusionnement et les conseils à donner aux superviseurs pour aider les différents types d'enseignants désillusionnés étaient différents pour les trois niveaux de maturation professionnelle.

Introduction

When completing a study to determine the supervision appropriate for the professional growth of teachers at three levels of development, an interesting and unanticipated result appeared (Sandell, 1990). Eighty percent of 42 elementary teachers interviewed felt that teachers go through a disillusionment phase and 69% of those teachers had personally experienced or were currently experiencing disillusionment with their profession.

The issue of teacher disillusionment is rarely addressed in the literature, although Trew-Williams (1985) described her fourth stage as a "soul searching, cynical phase." It came when teachers asked themselves, "What am I doing this for?" as they observed other teachers who put in the minimum amount of effort and had the same advantages as they did. The disillusionment was not related to their competence, but to their perception that the system did not appreciate their contribution or efforts.

At first, we took the position that although some kind of soul searching may be commonplace, the notion of a distinct disillusionment stage for all

Murray Sandell is the principal of Hants West High School in Nova Scotia. His research centers on the supervision of teaching.
Keith Sullivan is director of the School of Education. His research interests focus on the administration of educational institutions.

teachers was probably overstated as disillusionment might be confused with temporary dissatisfaction or disgruntlement. However, as the study progressed it became obvious that teachers' disillusionment with their profession was widespread and important.

Methodology

Because teachers so often reported the impact of supervision on professional development as inconsequential (McCarty, Kaufman, & Stafford, 1986; and Thompson & Zeimer, 1982), the original study (Sandell, 1990) was designed to examine the commonalities and differences with respect to their professional development, as well as appropriate supervision for teachers. Burden (1982) concluded that one reason supervision might be ineffective was the failure to provide different types of supervisory assistance for teachers at different developmental levels. Some adult and teacher development literature, for example Trew-Williams (1985), describe a differentiation with respect to the levels of professional development. From this literature and the experience of the authors, three levels of professional development were designed for analyzing elementary teachers as adult learners.

The label *professionally maturing* was used for the teachers at the highest of the three levels of professional growth in lieu of *renewal/mature*, often used in the literature, because the former suggests a continuous process of professional development, whereas the latter implies "having arrived." The middle level of professional development, *the consolidated learner*, came directly from the stage literature (Feiman-Nemser, 1983; Garman & Hazi, 1988; Huberman, 1988) and involved consolidated maintenance and embryonic consolidation, adjustment, integration, stabilization and maintenance of pedagogical theories and issues. The *embryonic teacher* described teachers at the lowest of the three levels of professional development and was not found in the literature. The term *embryonic* in this study implies a lack of professional growth, not the amount of teaching experience (Burden, 1982; Fuller & Bown, 1975). It is preferable to the often-used labels *novice* and *beginner*, which seem to indicate that the teachers are neophytes.

The primary source of data was a semistructured, in-depth interview of 30 to 40 minutes with each of 42 elementary teachers from a school system in Nova Scotia. The methodology was similar to Herzberg's (1966) *critical incidents* method used to determine the two-factor theory of motivation. This methodology requires each person to describe events, or critical incidents, personally experienced at work that produce a marked improvement or reduction in their job satisfaction.

The 42 participants were from a single district in Nova Scotia serving a mixture of rural and urban students and were selected by their principals to represent one of the three levels of professional development. Principals from 15 schools volunteered for the study. Only elementary schools were selected because of the documented professional development differences between secondary and elementary teachers (Cuban, 1983).

Using teachers from the same system opens the possibility that a major cause of disillusionment might be the district policies, which would make generalizations of the findings difficult. However, in our opinion, this particular district was similar to other Nova Scotia districts, and if any differences

were prevalent with supervision, it would be in favor of the district policies, which seemed to be slightly advanced in formative, positive supervision.

A teacher profile instrument developed for the study from the developmental stage and non-stage literature in both adult learning and teacher development was used in the selection. The instrument listed the distinguishing characteristics of the teachers in each of the three levels of professional development. Each principal used this instrument to select a teacher for each category and forwarded this list to the researcher without linking the teachers with the categories.

The teacher sample was composed of 15 professionally maturing, 15 consolidated/maintenance, and 12 embryonic teachers. One of the researchers visited the teachers in their schools or homes and interviewed and taped for approximately 30 to 45 minutes, using a semistructured probing, in-depth questionnaire approach, during April and May 1989. The tapes were transcribed and the data analyzed according to constructed themes and unexpected themes of apparent significance that emerged from the data. The interviewer was informed of the principal's categorization of the teachers only after the interviews, transcripts, and initial analysis had been completed.

Among the questions the 42 subjects were asked were "Do teachers go through a disillusionment stage?" "Have you gone through one?" "What, in your opinion, was the cause?" Later during the interview participants were asked, "What advice would you give to a supervisor working with a disillusioned teacher?"

Findings

Causes of Disillusionment

Of the 42 teachers in the sample, 69% indicated that they had experienced a major disillusionment phase in their career. Table 1 presents the reported causes of disillusionment.

Both the professionally maturing and consolidated/maintenance groups of teachers agreed on the major causes of disillusionment, that is, frustrations with relationships and lack of environmental stimulation. The teachers at the embryonic stage of development cited two different major causes of their disillusionment: bad classes, problem students, and no appreciation or support. It is interesting to note that the professionally maturing teachers cited no appreciation or support as the least important cause of their disillusionment. Whereas no embryonic teachers cited personal problems as a cause of their disillusionment, 33% of the professionally maturing teachers cited personal problems as a cause.

In many ways the causes of disillusionment reinforce the differences among the categories of teachers in this study, particularly between the professionally maturing and embryonic teachers. Some of the more specific comments that related to frustrations with relationships were the negative impact of staffroom conversations, superficiality of staff interactions, and professional development directed solely to raising one's teaching certificate, which at times distracted teachers from preparation and delivery of lessons, with enthusiastic teachers being ostracized by others more concerned with the routine day-to-day activities.

Table 1
Percent and Rank of Teacher Cited Causes of Their Own Disillusionment by Category of Professional Development

	PM		CM		E		Total	
	%	Rank	%	Rank	%	Rank	%	Rank
Lack of environmental stimulation	47	2	40	1*	25	4	38	2
Bad class, problem students	13	4	33	3	42	1*	29	3
Frustrations with relationships	67	1	40	1*	33	3	48	1
Personal problems	33	3	13	5	0	5	17	5
No appreciation or support	7	5	20	4	42	1*	21	4
Total Responses	25		22		17		64	

Note: Columns total more than the sample in each category because respondents could offer more than one response.
PM = Professionally Maturing (N = 15)
CM = Consolidated/Maintenance (N = 15)
E=Embryonic (N = 12)
*Tied scores

The no appreciation or support cause of disillusionment involved fear and confusion about the evaluation process as compared with the supervision process. The common "watchdog" focus and the alleged artificiality and formality of evaluation promoted disillusionment with the supervisors and the supervision process. With respect to artificiality, participants referred to the futility, superficiality, and formality of the routines for no substantive gains. The "watchdog" focus meant the control and inspection of teachers for the purpose of keeping teachers in line.

Environmental stimulation involved collaboration among teachers from the same staff or the same district as they traveled to and from courses and participated cooperatively in assignments. Such experiences appeared to combat boredom and disillusionment and caused stimulation for professional growth. Also teachers ranked courses, followed by inservice sharing with other teachers, informal acts of the principal, and conferences as the most important activities and experiences relating to environmental stimulation. Informal acts of the principal included showing interest, modeling appropriate behavior, being available to discuss problems, recognizing achievement, and encouraging experimentation.

Supervision

Table 2 presents the information from teachers of different developmental stages giving specific advice to supervisors for assisting disillusioned teachers.

The category of support, feedback, empathy, and understanding was the first or second most often mentioned advice by all three levels of teachers. The professionally maturing and consolidated/maintenance teachers included

Table 2
Percent of Teachers by Teacher Professional Development Stages and Advice to Supervisors Specific for Assisting Disillusioned Teachers

	PM		CM		E		Total	
	%	Rank	%	Rank	%	Rank	%	Rank
Support, feedback, empathy, understanding	47	2	73	1	75	1	64	1
Facilitate contact with professional development resources	53	1	33	2	16	3*	36	2
Informal talk, get to know teacher interests	33	3	7	5	16	3*	20	3*
Facilitate study break, leave	20	4*	13	3*	25	2	20	3*
Move, shift teacher	20	4*	13	3*	8	5	14	5
Total Responses	26		21		17			

Note: The total number of responses exceeds the sample size because respondents were encouraged to offer more than one response.
PM = Professionally Maturing (N = 15)
CM = Consolidated/Maintenance (N = 15)
E=Embryonic (N = 12)
*Tied scores

facilitating contact for teachers with professional development resources as important advice to supervisors. The second most often mentioned advice by the embryonic teachers was to facilitate study breaks or leaves.

Only 48% felt at this point in their careers that supervision had contributed positively to their professional development, whereas 41% believed that it had made no contribution. It appears that supervision, perceived by so many to be not helpful may have contributed to disillusionment.

In the study 71% of the respondents envisioned the supervision process as being "potentially" helpful, and they desired a supervisor who would accord them support, suggestions and feedback, and professional literature. An equal number preferred a model that permitted collaboration with peers and constructive criticism by the supervisor, who would also work with students in the classroom.

Most respondents (81.5%) believed that teachers should have a more or less equal role in the supervision process, while 74.2% favored the notion of teachers supervising teachers, noting the collegial relationship they already enjoyed and the expertise of teachers compared with that of principals. Others (22.8%) who opposed the concept believed that their relationship with peers would suffer, that colleagues would be overly critical, or that they could not bring themselves to criticize their peers or to accept criticism from them. Some pointed to the need for an authority figure. Supporters and detractors alike agreed on the need for the training of supervisors.

Most participants, responding to some direct questions with respect to teacher supervision, reported that they had been supervised once every three

years and described a traditional checklist form of evaluation. Many expressed anxiety with regard to this process. However, the majority (60.5%) perceived that the supervision process had changed positively over the last few years.

The artificiality, formality, judgmental, and threatening natures of supervision were considered to be major barriers to effective professional development, followed by infrequency of supervision, the incompetence of the supervisor, and the distance between supervisor and teacher. All these factors seemed to contribute to disillusionment.

The most common advice of the respondents for the functioning of the supervisor, was to establish a positive supervisor-teacher rapport and to get to know their teachers, their interests, expertise, philosophy, and stage of development. Next they stressed humanizing the supervision process, that is, showing kindness, encouraging teacher professional development, and focusing on their own professional and personal development. This advice is similar to that already discussed to assist a disillusioned teacher.

The overwhelming majority (95%) of teachers claimed their growth came from informal influences, including informal interactions with fellow teachers and the informal acts of principals, described as showing interest, modeling appropriate behavior, being available to discuss problems, recognizing achievement, and encouraging experimentation. Informal influences such as discussions about the practice of teaching, opportunities to observe another's teaching, and to teach and learn from one another were deemed more significant than formal influences, that is, conferences, inservices, supervision and courses. The informal influences took place in training sessions, faculty meetings, teachers' lounges, hallways, and classrooms.

Ninety-two percent pointed to the informal, everyday acts of the principal as important to teacher professional development, including disseminating professional materials, focusing professional interests around specific themes, soliciting teacher responses to educational concerns, encouraging experimentation and motivation, and publicizing achievements. Teachers felt that growing professionally through informal interaction with their peers or from the informal everyday acts of the principal, they would be less prone to fall prey to disillusionment. Such informal influences would provide relationships, stimulation, appreciation, support, advice, and possibly a "listening ear" for personal problems.

It is interesting that teachers would perceive supervisors as being personal counselors. This perception seemed to be based on supervisory activities that included modeling appropriate attitudes, providing support by encouraging experimentation, publicizing achievements, and listening to teachers' problems.

Conclusions

Supervisors should attend to the causes of disillusionment among their teachers by stages of professional development, first, because it is a prevalent condition, second, because they can play an important role in preventing disillusionment, and, third, because they can assist teachers to see their profession in a positive light.

Teachers at the middle level of professional development (consolidated learner) and teachers at the highest level of professional growth (professional-

ly maturing) both were most disillusioned by frustrations with relationships and lack of environmental stimulation. In more specific terms, their frustrations were with staffroom conversations, superficiality of interactions with other staff members, colleagues being solely interested in professional development to raise one's certification for monetary reasons, and enthusiastic teachers being ostracized by others more concerned with routine activities. The teachers who were not showing professional growth (embryonic) most often cited bad classes, problem students, and no appreciation or support as their cause of disillusionment. In addition, one third of the teachers at the highest level of professional development indicated that personal problems added to their disillusionment, whereas none of the embryonic teachers perceived personal problems as a cause of their disillusionment.

These results indicate that supervisors cannot treat all teachers equally when attempting to bring teachers out of a disillusionment phase. The most professional teachers also need assistance with their personal problems, better relationships with their colleagues and more environmental stimulation. With these teachers, supervisors should encourage professional development by joint course enrollments, personal and informal acts of the principal, and involvement in conferences. Teachers who are least professionally developed need more direct help with their classes and problem students, along with greater appreciation and support for their work. They appeared to be fearful and confused about the evaluation process, mainly seeing it as artificial and summative.

When asked to provide advice to supervisors, teachers at all levels of professional development mentioned supportive feedback and empathy and understanding as the first or second most important advice to supervisors for helping disillusioned teachers. The two highest levels of professionally developed teachers included facilitating contact for teachers with professional development resources as important advice to supervisors, while the embryonic teachers' second most often mentioned advice was to facilitate study breaks or leaves.

A significant number of professionally maturing teachers had other advice to offer, such as informal discussions for getting to know teachers' interests, study breaks, and moving teachers, whereas the consolidated/maintenance teachers rarely gave informal discussions and getting to know teachers' interests as important advice. The embryonic teachers rarely advised moving teachers from one class situation or school to another.

Overall, over 40% of the teachers felt that supervision had not contributed positively to their professional development to this point in their career, yet at the same time almost three quarters of the teachers felt that supervision had potential for being helpful, particularly to the disillusioned teacher. Disillusionment with teaching in general, and supervision specifically, may be related to supervisors not treating teachers at different levels of professional maturity in the same ways.

This study highlights the fact that teachers are at different stages of professional development, that they will almost all go through a stage of disillusionment with their profession, and that teacher disillusionment must be treated differently depending on the stage of professional maturity. Proper teacher

supervision based on stages of professional maturity could advance professional development that would simultaneously combat teacher disillusionment.

References

Burden, P.R. (1982). *Developmental supervision: Reducing teacher stress at different career stages.* Paper presented at the Annual Meeting of the Association of Teacher Educators, Phoenix. (ERIC Reproduction Service No. 28267)

Cuban, L. (1983). Effective schools: A friendly but cautionary note. *Phi Delta Kappan, 64,* 695-696.

Feiman-Nemser, S. (1983, June). *Learning to teach.* Washington, DC: National Institute of Education.

Fuller, F., & Bown, O.H. (1975). Becoming a teacher. In K. Ryan (Ed.), *Teacher Education The seventy-fourth yearbook of the National Society for the Study of Education* (pp. 25-51). Chicago IL: University of Chicago Press.

Garman, N.B., & Hazi. (1988). Teachers ask is there life after Madeline Hunter? *Phi Delta Kappan, 69,* 669-72.

Herzberg, F. (1966). *Work and the nature of the man.* Cleveland, OH: World Publishing.

Huberman, M. (1988). Teacher careers and social improvement. *Journal Curriculum Studies, 20*(2), 119-32.

Katz, L.G. (1985). Research currents: Teachers as learners. *Language Arts, 62*(7), 778-82.

McCarty, D.J., Kaufman, J.W., & Stafford, J.C. (1986, April). Supervision and evaluation: Two irreconcilable processes? *The Clearing House,* 351-352.

Sandell, M.W. (1990). *The supervision appropriate for elementary teachers as adult learners at various levels of professional development,* Unpublished doctoral dissertation, Dalhousie University.

Thompson, A.G., & Zeimer, R.M. (1982, May). Supervision of instruction: Dying or dead? *The Clearing House, 55*(9), 394-96.

Trew-Williams, L.D. (1985). *Developmental patterns of teaching careers.* Published doctoral dissertation, University of Alberta.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 2, June 1992, 141-153*

Reg Fleming
University of Saskatchewan

Teachers' Views of Technology

This article reports the results of a survey of 1,200 teachers' views on technology. Specifically, teachers' responses to existing statements about technology in contemporary society were examined. As well, the paper describes new statements about technology based in the sociology of technology and describes teachers' responses to these statements. A series of statement pairs (positive and negative exemplars) were designed to elicit teachers' responses to the issues of technology and quality of life, technology and employment, technology and health care, technology and food production, technology and social well-being, and technology and its social control. The responses indicated that the teachers in this sample overwhelmingly described and critiqued technology from what is described in the paper as an artifact perspective, that the majority hold an incomplete view of the nature of technology, and that most feel technology is beyond the control of the individual. No significant relationship was found between these responses and sex, age, teaching experience, subjects taught, school size, and preservice education. Implications for current classroom practice and preservice education are discussed.

Cet article présente les résultats d'une enquête portant sur les vues de 1200 enseignants concernant la technologie. Plus précisément, on a analysé les réponses des enseignants à des questions d'actualité au sujet de la technologie dans la société contemporaine. L'article décrit également de nouveaux points de vue fondés sur la sociologie de la technologie, et les réactions des enseignants à ces points de vue. Une série de paires de déclarations (dont l'une positive et l'autre négative) était concue pour obtenir les réponses des enseignants aux questions de technologie et qualité de la vie, technologie et emploi, technologie et santé, technologie et production alimentaire, technologie et bien-être social, et technologie et contrôle social. Les réponses ont montré que la plupart des enseignants de cet échantillon ont décrit et critiqué la technologie à partir de ce qu'on appelle dans cet article une perspective d'artefact; que la majorité a une vue incomplète de la nature de la technologie, et que la plupart trouvent que la technologie est au-delà du contrôle de l'individu. Il n'y avait pas de relations significatives entre ces réponses et le sexe, l'âge, l'expérience dans l'enseignement, les matières enseignées, la grandeur de l'école, et la formation initiale. Les implications pour la pratique éducative en cours et la formation initiale sont présentées.

Introduction

In several Canadian provinces (British Columbia, Alberta, Saskatchewan) a number of curricular changes, particularly in the areas of science and social studies, have led to an increased emphasis on the study of technology, with the hope of creating technologically literate school graduates.

Describing technology and technological literacy has turned out to be a complex task. A comprehensive overview of the issues and problems encountered when attempting this description was the focus of a double issue of the *Bulletin of Science, Technology, and Society* (1986). A close examination of this issue indicates that the task was not totally accomplished. Staudenmaier

Reg Fleming is an associate professor in the Department of Curriculum Studies. His interests include science, technology, and society education, curriculum theory, and qualitative research issues.

(1985) offers a thorough analysis of the perspectives on technology offered in all pertinent papers in the journal *Technology and Culture*. Although these perspectives are of immense help to the theoretician, they were never written with the educational researcher in mind.

Fleming (1989) offered a combination of a theoretical model of technology with its concomitant instructional implications. A portion of the model is based in Kline (1985). He starts with common usage, namely, that technology is hardware. By hardware he means all non-natural objects manufactured by humans. He then describes the second most common usage, that of technology as the process of manufacturing the hardware. Kline argues that this descriptor must include all elements necessary to manufacture particular hardware. These would include people, machines, resources, as well as the physical, legal, economic, and political environments. He labels this the sociotechnical system of manufacture.

The third perspective suggested by Kline (1985) is that of technology as know-how. Know-how is the information, skills, processes, and procedures for accomplishing tasks. These three usages, he claims, form the common view of technology.

Kline (1985) goes on to argue that these three usages are inadequate. He suggests that a fourth usage is necessary to give purpose to the manufacture of hardware. This fourth concept he labels the sociotechnical system of use. It contains the hardware and the people necessary to achieve the purpose of extending human capacities. As an example, says Kline:

> We embody automobiles in a system of roads, gas stations, laws for ownership and operation, rules of the road, and use the combined system (the autos plus all the rest) to extend the human capacity for moving ourselves and our possessions about. (p. 216)

Without a sociotechnical system of use, the manufacture of hardware has no purpose. Kline (1985) elaborates further by stating that "sociotechnical systems of manufacture and sociotechnical systems of use form the physical bases of all human societies past and present" (p. 217). This pattern of purposeful innovation in sociotechnical systems distinguishes humans from other animals. Kline concludes by saying that "few topics are more basic ... than an understanding of the nature of sociotechnical systems and the pattern in which we humans use them to create the physical bases for our societies past and present" (p. 218). This description of technology, which combines the four perspectives described above, is referred to as sociotechnology. A person understanding sociotechnology would be technologically literate because this person has the power and the freedom to use that power to examine and question issues raised by technology. This empowerment would enable such a person to become critical about technology. Being critical must never be confused with being antitechnology. Rather, being critical about technology means having the intellectual skills to examine the pros and cons of any technological development, to examine its potential benefits, its potential costs, and to perceive the underlying political and social forces driving the development. It was this model of sociotechnology that served as the theoretical base for several aspects of the research described in this article.

A significant volume of research does exist that describes students' views on technology. The views of high school graduates (Aikenhead, Fleming & Ryan, 1987; Aikenhead & Ryan, 1987) and the views of undergraduate science students (Fleming, 1988) have been explored. However, very few research studies are designed specifically to document teachers' views on technology. Lack of data in this area poses serious problems for those who wish, for example, to design preservice and inservice programs for those teachers who must implement curricula with a sociotechnology focus. In Saskatchewan, recent curriculum reform has resulted in the creation of six Common Essential Learnings (CELs), one of which is called technological literacy. It reflects the concept of sociotechnology described above.

These CELs, where possible, are to be presented in all core subjects through all grades. Technological literacy, then, is to be a strand woven through all subjects. This is a completely new curriculum emphasis that is not reflected in most existing materials. Given that in teachers' academic and professional education little or no time has been spent on the study of sociotechnology, the teacher response has been as expected. Requests for inservice programs are constant. As previously mentioned, lack of knowledge about teachers' views has become a stumbling block. It was this lack of knowledge that prompted the research described below.

Purpose

The purpose of this research was to conduct an exploratory study of teachers' views on the nature of technology.

Design and Procedures

This study used survey research methods. A random sample of 1,200 teachers was drawn from the entire K-12 teaching population (N=12,000) of the province of Saskatchewan. The list of all teachers was provided by the Saskatchewan Teachers' Federation. The research staff of the Federation drew the sample. Each teacher in the sample was mailed the survey instrument (described below). After 10 days, a reminder card was sent. Five hundred, ninety-six teachers (49.6%) completed and returned the survey.

The Survey Instrument

The instrument was comprised of three sections. The first section asked teachers to respond to statements about technology and its relationship with science taken from the instrument *Views on Science-Technology-Society* (VOSTS) form mc.3 (Aikenhead & Ryan, 1987). These are listed in Appendix A. The response format was closed, that is, for each VOSTS statement, a range of responses was offered to the respondents. These responses represent categories that were generated in earlier work (Aikenhead, Fleming, & Ryan, 1987) by analyzing written responses to these statements. The possible responses are called Teacher Position.

The second section of the instrument asked teachers to offer short written responses to 14 statements created by the researcher about technology arising from the literature in the sociology of technology. These statements dealt with the relationship between technology and employment, the nature of technological development, personal standard of living and technology, the relationship between technology and social change, technological determinism,

and specific technological developments. The format of the statements and the methods for testing their reliability was based on earlier reported work (Aikenhead, Fleming, & Ryan, 1987). The teachers in the sample were randomly assigned to four subgroups (N=300) for this section. Each subgroup received four of the new statements. These statement subgroupings are presented in Appendix B. (Note that two of the new statements were repeated. These were randomly selected.) A response rate of 50% ensured a theoretical maximum of 150 paragraph responses for each statement. The nonusable responses are tabulated in the analysis section and were so small (1-3%) as to guarantee a large number of usable paragraphs for analysis for each statement. The research cited earlier had found for the analysis of written argumentative responses theoretical saturation was reached by or prior to the 30th response.

The third section of the instrument asked for demographic information.

Analytical Procedures

The VOSTS results were analyzed using SPSS-X to generate response profiles for each statement on the instrument. The written responses were analyzed according to the method of Aikenhead, Fleming, and Ryan (1987). For each of the new statements, the written paragraph responses were analyzed to discern common arguments that, when categorized, represent teachers' common viewpoints. These categorized viewpoints become the teachers' positions reported in this paper.

Combining the demographic data and the response profiles, Pearson correlation coefficients were calculated to determine whether there were relation-

Table 1
Demographics of the Respondents

Gender		*Median Age*	
Female	58%	38 years	
Male	42%		
		Median years of teaching experience	
		13 years	
*Division Mainly Taught**		*Student Population of School*	
Division I	27%	Fewer than 100	9%
Division II	22%	100 - 250	38%
Division III	15%	250 - 500	33%
Division IV	30%	500 - 750	11%
Other	6%	750 - 1000	5%
		1000 - 1250	2%
		More than 1250	2%
University courses in any physical sciences			
No	67%		
Yes	33%		

*Division I (Grades K-3), Division II (Grades 4-6), Division III (Grades 7-9), Division IV (Grades 10-12)

ships between VOSTS responses and sex, age, years of teaching experience, subjects taught, school size, and preservice science education.

Results

The results are presented in four sections. The first summarizes the demographic data, the second presents the VOSTS results, the third presents the results of the analysis of the argument responses to the new statements, and the fourth the analysis of the VOSTS/demographics interactions.

The Demographic Data

The demographic data are presented in Table 1. As the reader is well aware, attention must be paid to the representativeness of the respondents vis-à-vis the population. The population data for teachers show a median age of 38 years, with the median years of teaching experience at 13 years. This is identical to the figures for the respondents. As well, the gender percentages and student population percentages closely reflect the population parameters. Nevertheless, the response rate is such that one must be cautious in generalizing to the entire teaching population.

Table 2

VOSTS 12.1	In order to improve the quality of living in Canada, it would be better to invest money in technological research RATHER THAN scientific research.	
	Teacher position	*% of Usable Responses*
A.	Invest in *technological* research because it will improve production, economic growth, and unemployment. These are far more important than anything that scientific research has to offer.	1
B.	There is really no difference between science and technology.	0
C.	Scientific knowledge is needed to make technological advances.	10
D.	They interact and complement each other equally. Technology gives as much to science as science gives to technology.	26
E.	Each in its own way brings advantages to society. For example, science brings medical and environmental advances, while technology brings improved conveniences and efficiency.	46
F.	Invest in *scientific* research—that is, medical or environmental research—because these are more important than making better appliances, computers or other products of technological growth.	3
G.	Invest in *scientific* research because it improves the quality of life (e.g., medical cures, answers to pollution, and increased knowledge). Technological research, on the other hand, has worsened the quality of life (e.g., social welfare, education, job creation programs, the fine arts, foreign aid, etc.)	2
H.	Invest in *neither*. The quality of living will not improve with advances in science and technology, but will improve with investments in other sectors of society (e.g., social welfare, education, job creation programs, the fine arts, foreign aid, etc.)	3
I.	I don't understand.	0
J.	I don't know enough about this subject to make a choice.	4
K.	None of these choices fit my basic viewpoint.	5

Results of the Analysis of the VOSTS Statements

A general comment is in order first. The VOSTS item numbers may appear to be erratic. The reader is reminded that these are the VOSTS statements which deal with technology and are drawn from a larger, more orderly collection of statements.

1. *Science, technology, and the quality of life: VOSTS 12.1 and 6.1.*

The data for VOSTS 12.1 are presented in Table 2. VOSTS 12.1, a general statement concerning quality of living and science and technology requires that the respondent differentiate between science and technology and use this differentiation when responding. Seventy-two percent of the responses cluster around two items, D and E. Position D, representing 26% of the responses, offers a contemporary view of the nature of the relationship between science and technology, namely, that they interact and complement each other. Position E, representing 46% of the responses, posits a functionalist role for science in which science directly benefits society through medical and environmental advances. This suggests a confusion over the social roles of science and technology, a suggestion further supported by the responses to the next item, VOSTS 6.1.

VOSTS 6.1 moves from the more general statement, quality of living, raised in 12.1 to specific social problems such as poverty, crime, unemployment, overpopulation and the threat of war. The data from VOSTS 6.1 are presented in Table 3.

Once again, a clear understanding of the nature of science and the nature of technology is required. As can be seen from Table 3, half the respondents offer a populist view of science and technology in which these enterprises are under

Table 3

VOSTS 6.1	Science and technology offer a great deal of help in resolving such social problems as poverty, crime, unemployment, over-population, and the threat of a nuclear war.	
	Teacher position	*% of Usable Responses*
A.	Science and technology can certainly help to resolve these problems. The problems could use new ideas from science and new inventions from technology.	4
B.	Science and technology can help resolve some social problems but not others.	6
C.	Science and technology solve many social problems, but science and technology also *cause* many of these problems.	29
D.	It's not a question of science and technology helping, but rather it's a question of *people* using science and technology wisely.	50
E.	It's hard to see how science and technology could help very much in resolving these social problems. Social problems concern human nature; these problems have little to do with science and technology.	6
F.	Science and technology only make social problems worse. It's the price we pay for advances in science and technology.	1
G.	I don't understand.	0
H.	I don't know enough about this subject to make a choice.	1
I.	None of these choices fit my basic viewpoint.	3

Table 4

VOSTS 11.2	Defining what technology is can cause difficulties because technology does many things in Canada. But MAINLY technology is:	
	Teacher Positions	*% of Usable Responses*
A.	Very similar to science.	0
B.	The application of science.	26
C.	New processes, instruments, tools, machinery, appliances, gadgets, computers, or practical devices for everyday use.	26
D.	Bombs, military hardware, nuclear reactors, etc.	0
E.	Robotics, electronics, computers, communication systems, automation, etc.	4
F.	A technique for doing things.	5
G.	A way of solving practical problems.	2
H.	Inventing, designing and testing things (e.g., artificial hearts, computers, space vehicles).	4
I.	Ideas and techniques for designing and manufacturing things, for organizing workers, business people and consumers, for the beneficial progress of society.	27
J.	I don't understand.	0
K.	I don't know enough about this subject to make a choice.	1
L.	None of these choices fit my basic viewpoint.	6

the wise control of people (Position D). Given the exploratory nature of this study, interviews with selected respondents could clarify the meaning of "people." Which people the respondents refer to should be explored further. Twenty nine percent of respondents (Position C) confuse the role of science and technology. It appears that, as in previous research findings (Fleming, 1987a), they lump the two terms together as "technoscience" and see no differing roles for the two. This confusion can be examined in the light of the data in the next two tables.

2. *The relationship between science and technology: VOSTS 11.2, 11.3 and 39.*

VOSTS 11.2 (Table 4) offers insight into the practical definition(s) of technology offered by the respondents.

For all intents and purposes, there are three positions. The first, item B, views technology as applied science. It has been argued elsewhere (Barnes & Edge, 1982, p. 140) that this is an untenable position. The second, item C, sees technology as know-how and artifacts. This restricted meaning for technology (Fleming, 1987b; 1989) will be seen again in the analysis of arguments under the heading Results of the Analysis. In all, 52% of the respondents hold an incorrect or restricted view of technology. The final group (27%) opt for item I, a position much closer to the view of technology called sociotechnology, in which technology is seen as a social process.

VOSTS 11.3 (Table 5) explores the nature of the relationship between science and technology. As argued earlier (Fleming, 1987a), unless specifically asked to do so, people do not differentiate between the two enterprises. This was the case in VOSTS 12.1. where "technoscience" best describes the respondents' views of science and technology. This item had the largest

Table 5

VOSTS 11.3	Science and technology are closely related to each other.	
	Teacher Position	*% of Usable Responses*
A.	Science is the basis of all technological advances; though it's hard to see how technology could aid science.	1
B.	Scientific research leads to practical applications in technology, and technological developments increase the ability to do scientific research.	87
C.	Although they are different, they are linked so closely that it's hard to see how science could aid technology.	4
D.	Technology is the basis of all scientific advances; though it's hard to see how science could aid technology.	1
E.	Science and technology are essentially the same thing.	1
F.	I don't understand.	0
G.	I don't know enough about this subject to make a choice.	2
H.	None of these choices fit my basic viewpoint.	3

proportion of answers. As can be seen in Table 5, 87% of the sample chose item B, which could present a reciprocity between science and technology. This is a marked increase from the 26% in VOSTS 12.1 who chose position D, a comparable response.

VOSTS 39 explores an issue that is very evident in Canadian media, namely research and development (R & D). Ziman (1984) offers the following insight into R & D:

> Like the distinction between science and technology, the distinction between "research" and "development" is becoming more and more difficult to discern. The combination "Research and Development"—"R & D"—can be used to cover a whole range of investigatory procedures, from the most abstract theoretical analysis to the most earthly trial and error, which may be found in all fields of "S & T," whether motivated by a desire for fundamental understanding, as in cosmology, or by a desire to make a commercial profit, as in the manufacture of weapon systems. (p. 122)

Examining VOSTS 39 results (Table 6), one finds a wider range of responses to this item. Items C and D constitute the major (54%) response. In both of these items, the description of research is constant. The description of development changes. In item C, development is a functionalist process that benefits society. In item D, development is much more closely related to Ziman's (1984) strategic research. According to Ziman, "Strategic research is also knowledge-oriented but it has a conscious utilitarian purpose, such as the determination of the facts and theoretical principles relevant to existing or projected technologies" (p. 142). The 13% of respondents who chose item F seem closest to a contemporary but simplistic view of R & D. Item H reflects the concern over the use of R & D.

Results of the Analysis of the Argument Responses

As mentioned earlier, 14 new statements based in the sociology of technology were used with sample subgroups. These are found in part two of the questionnaire (Appendix B). The format of the statements and the analysis of the

Table 6

VOSTS 39	Science and Technology are important to the research and development (R & D) in Canadian industry. What does research and development mean to you?	
	Teacher Position	*% of Usable Responses*
A.	R & D means finding new answers to questions about the world and about people.	3
B.	R & D means progress by making life easier and the quality of life better.	2
C.	Research is exploring for new facts, ideas and information. Development is putting them to use in order to benefit society.	35
D.	Research is exploring for new facts, ideas and information. Development is putting them to use by coming up with new and creative ideas.	19
E.	R & D means exploring new ideas and problems in industry, in order to help an industry overcome its problems and thereby produce newer and better products.	5
F.	R & D means a combination of science and technology. Research leads to development, and development leads to improved research.	13
G.	R & D means helping humanity by finding medical cures and new technologies. But *unanticipated* effects of R & D can also cause social problems.	5
H.	R & D usually means helping humanity by finding medical cures and new technologies. But R & D also means harming society by creating such things as nuclear arms and other wasteful or unhealthy technologies. It depends on the R & D or how it is used.	17
I.	I don't understand.	0
J.	I don't know enough about this subject to make a choice.	1
K.	None of these choices fit my basic viewpoint.	2

written arguments has been described earlier (Aikenhead, Fleming, & Ryan, 1987).

As described in this paper, participants were asked to respond with "agree," "disagree," or "can't tell" to their statements. As well, they were to offer a justification for their choice in the form of an argumentative paragraph. The central focus of the analysis is the paragraph. For each paragraph, I wrote a phrase that summarized the argument. The collected phrases were examined to determine common arguments. These became the "teacher positions." A general comment is in order. The overwhelming response to most of the statements was to reason with *artifacts*. Simply put, for many, in response to statements B-1, B-2.1, B-2.2, B-4.1, B-4.2, B-5.1, B-5.2, B-6.1, B-6.2, the name of an artifact—microwave oven, VCR, satellite TV—was used as a replacement for a reasoned argument.

For example, the responses to the statement B-4.1 "Because of technology, my standard of living continues to improve" fell into two camps. One minority (15%) chose to offer arguments based on improved quality of life. A second minority (5%) argued it was one's finances, rather than technology, that determines one's standard of living. The majority (80%) responded with the name of an artifact. The arguments took the form of a single sentence

Table 7

Technological changes result in social changes (B-7.1).

	Teacher Position	*% of Usable Responses*
A.	My lifestyle relies on technology (16 artifacts are names, including the computer, word processor, automated tellers, and computerized axle flow combines)	83
B.	There is an improved quality of life (leisure time).	4
C.	Demographic patterns change.	9
D.	The structure of the family unit has changed.	3
E.	Non-usable.	1

response: "Yes it has, think of the ——," where the blank can be filled in with a specific artifact. For this item 32 artifacts were named.

When subjects were asked to respond to statements about specific technologies—medical technology (B-5.1/5.2) and agricultural technology (B-6.1/6.2) in this case—the "reasoning by artifact" process again dominated, with over 90% of the respondents naming specific medical or agricultural artifacts. Items B-7 and B-8, dealing with the relationship between technology and social change, also follow this pattern.

The following two tables, Tables 7 and 8, present the data offered in response to items B-7.1 and B-7.2. As can be seen, 83% of respondents offered the reasoning-by-artifact responses described above.

In Table 8, the issue of social change leading to technological change is raised. The majority position (Item A, 77%) is that society has "needs and demands" that force a direction on technological development. A close examination of these needs suggests that many of them are artifacts. One may infer that social change is seen as the result of or need for artifacts.

The final items, B-8.1 and B-8.2, are concerned with the control of technology. Responses to item B-8.2 are concerned with the control of technology. In response to item B-8.1, "Technological developments are beyond the control of the average citizen," 75% of respondents felt that average citizens have no control over technology, mostly because they lack knowledge of or the desire to learn about technological development. The remaining 25% felt that citizens

Table 8

Social changes result in technological changes.

	Teacher Position	*% of Usable Responses*
A.	Society's needs and demands dictate the direction technology takes. (Note: 12 needs and demands were identified, including curing illness, mass transportation, home security systems, and time-saving devices).	77
B.	The two are interdependent.	7
C.	I believe the reverse is true.	12
D.	Social change doesn't have to result in technological change.	1
E.	Non-usable responses.	3

could control technology through group actions such as public meetings and demonstrations.

In response to item B-8.2, "The average citizen can help control technological developments," a smaller number (16%) felt that the average citizen was not well enough informed about technology. The remaining respondents were split into three camps. One group (38%) felt that citizens had no control over development, but as consumers had control over the use of the development. Once again, many artifact examples were offered. A second group (26%) felt that we were helpless as individuals, but as a group we could possibly have an impact. The group activities suggested were again meetings and demonstrations.

The third group (18%) argued for controlling technological development through exercising democratic rights. These included voting and writing elected representatives. The remaining 2% felt that development was beyond citizen control because technology was controlled by large corporations and governments, neither of which was responsive to citizen input.

The VOSTS/Demographic Interactions

There were no statistically significant relationships found between VOSTS responses and sex, age, years of teaching experience, subjects taught, school size, and preservice science education.

Discussion

The data strongly suggest that the teachers in this study have a restricted view of technology. Specifically, the majority see technology as applied science or as the products of technology (artifacts). One hastens to add that this consistent view of technology is not the teachers' fault. It is possible, for example, that preservice education presents and/or reinforces their view. A content analysis of preservice education would be in order. One does suspect, though, that such an analysis would show very little attention paid to the study of technology.

This leads one to speculate on the possible sources for such a consistent view. One must ask what sources constantly present technology as applied science and/or artifacts.

In analyzing student beliefs about STS topics, Aikenhead (1988) found that "their sources of ideas about scientists and the social and technological context of science, came from the mass media (TV, films, newspapers, magazines, and books)" (p. 613). He goes to argue that "one can hardly avoid the speculation that students *and* textbook writers acquire their image of scientific research from the same source—television" (p. 614). If one extrapolates from these ideas, it is quite reasonable to assume that most teachers receive their major impressions about science and technology from the media. To verify this, a detailed media analysis would be required, something similar to that done by Nelkin (1987) on print media.

For the immediate future, this study has several implications. First, in choosing materials for classroom use, teachers will choose those materials that most closely echo their own understanding. If the understanding is restricted, so will be the materials chosen. Second, these results suggest that most teachers will teach students either that technology is applied science *or* that technol-

ogy is artifacts. This will cause serious difficulties for those school jurisdictions attempting to implement new curricula with a technological literacy strand through all subjects.

What are implications for preservice education? I do not want to suggest that more coursework is the answer. Nevertheless, it might be appropriate to mandate courses in the sociology of science and of technology for prospective educators. If nothing else, such courses could help prospective teachers come to grips with contemporary arguments about the nature of these enterprises. Teaching methods that allow for a study of technology in all courses must also become a part of the prospective teacher's repertoire. Without these actions, kindergarten to grade 12 students will continually be offered restricted meanings for technology.

From a researcher's perspective, further exploration of these views should be carried out. Interviews with teachers to clarify the reasoning behind their views is essential. The data reported in this paper offer a starting point for these interviews.

Acknowledgement

This project was funded by the research branch of the Saskatchewan School Trustees' Association (SSTA). The views expressed herein are those of the author and do not necessarily represent the views of the SSTA.

References

Aikenhead, G.S. (1988). An analysis of four ways of assessing student beliefs about STS topics. *Journal of Research in Science Teaching, 25*(8), 607-629..

Aikenhead, G.S., Fleming, R.W., & Ryan, A.G. (1987). High school graduates' beliefs about science-technology-society. I. Methods and issues in monitoring students' views. *Science Education, 71*(2), 145-161.

Aikenhead, G.S., & Ryan, A.G. (1987). *High school students' views on science, technology, and society.* Ottawa: Social Science and Humanities Research Council Research Report #410-85-1267.

Barnes, B., & Edge, D. (Eds.). (1982). *Science in context.* Cambridge, MA: MIT Press.

Bulletin of Science, Technology, and Society. (1986). Technological literacy, 6, (2 & 3), 141-336.

Fleming, R.W. (1987a). High school graduates' beliefs about science-technology-society. II. The interaction of science, technology, and society. *Science Education, 71*(2), 163-186.

Fleming, R.W. (1987b). *Technological literacy.* Regina, SK: Saskatchewan Education.

Fleming, R.W. (1988). Undergraduate science students' views on the relationship between science, technology, and society. *International Journal of Science Education, 10*(4), 449-463.

Fleming, R.W. (1989). Literacy for a technological age. *Science Education, 73*(4), 391-404.

Kline, S.J. (1985). What is technology? *Bulletin of Science, Technology, and Society, 5*, 215-218.

Nelkin, D. (1987). *Selling science.* New York: Freeman.

Staudenmaier, J.M. (1985). *Technology's storytellers: Reweaving the human fabric.* Cambridge, MA: MIT Press.

Ziman, J. (1984). *An introduction to science studies.* Cambridge, UK: Cambridge University Press.

Appendix A

VOSTS form mc.3 Items on Technology

VOSTS 12.1	In order to improve the quality of living in Canada, it would be better to invest money in technological research RATHER THAN scientific research.
VOSTS 6.1	Science and technology offer a great deal of help in resolving such social problems as poverty, crime, unemployment, over-population, and the threat of a nuclear war.
VOSTS 11.2	Defining what technology is, can cause difficulties because technology does many things in Canada.

VOSTS 11.3	Science and technology are closely related to each other.
VOSTS 39	Science and technology are important to the research and development (R & D) in Canadian industry. What does research and development mean to you?

Appendix B

Newly Created Items on Technology

B-1	A country's well-being relies heavily on investment in high technology.
B-2.1	Most employment opportunities in the next 15 years will arise in high technology areas.
B-2.2	Most employment opportunities in the next 15 years will not arise from high technology.
B-3	Technological development occurs when a more sophisticated machine replaces a simpler one.
B-4.1	Because of technology, my standard of living continues to improve.
B-4.2	Because of technology, my standard of living is *not* improving.
B-5.1	Advances in medical technology have improved the quality of health care.
B-5.2	Medical technology has not improved the quality of health care.
B-6.1	Advances in agricultural technology have *not* improved farming.
B-7.1	Technological changes result in social changes.
B-7.2	Social changes result in technological changes.
B-8.1	Technological developments are beyond the control of the average citizen.
B-8.2	The average citizen can help control technological developments.

If undelivered return to:

The Alberta Journal of Educational Research
4-116 Education North
University of Alberta
Edmonton, AB
T6G 2G5

Publications Mail Registration Number 1436



THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XXXVIII NUMBER 3 SEPTEMBER 1992

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

The Alberta Journal of Educational Research

Published in March, June, September, and December
by the Faculty of Education, University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation, and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Editor: Robert H. Short

Consulting Editors

James Chapman
Massey University

Rod Clifton
University of Manitoba

Dianne Common
Lakehead University

Carol Crealock
University of Western Ontario

John Hattie
University of Western Australia

Billie Housego
University of British Columbia

Linda Larocque
University of Alberta

Stewart J.H. McCann
University College of Cape Breton

Les McLean
Ontario Institute for Studies in Education

Linda Phillips
Memorial University of Newfoundland

Sharon Robertson
University of Calgary

Faculty Advisory Committee

John G. Paterson (Chair)
Fern Snart
Joyce Edwards
Graham Fishburne

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada, from the Alberta Advisory Committee for Educational Studies, and from the Office of Research Services, University of Alberta.

The subscription rate is $32.00 per year for individuals, $40.00 per year for institutions. Add $8.00 for delivery outside Canada. Single copies are $8.00 each. Subscriptions and sales in Canada will be charged 7% GST. Please make cheques payable to *The Alberta Journal of Educational Research*. Back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.

Publications Mail Registration Number 1436

Faculty of Education
University of Alberta

The Alberta Journal of Educational Research

Volume XXXVIII, Number 3, September 1992

Stephen P. Norris	**155**	A Demonstration of the Use of Verbal Reports of Thinking in Multiple-Choice Critical Thinking Test Design
Judith K. Bernhard	**177**	Gender-Related Attitudes and the Development of Computer Skills: A Preschool Intervention
Ellen M. Regan *John A. Ross*	**189**	Professional Development for School Consultants: An Evaluation
Marziya Yasmin *Baha Abu-Laban*	**205**	Ethnicity and Occupational Inequality: A Reconsideration
Estelle S. Gellman *Mina Berkowitz*	**219**	Factors Perceived as Important in Teacher Evaluation
Dushyanta Persaud *Paul R. Madak*	**235**	Graduates and Dropouts: Comparing Perceptions of Self, Family, and School Supports

ISSN 0002-4805

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 3, September 1992, 155-176*

Stephen P. Norris
Memorial University of Newfoundland

A Demonstration of the Use of Verbal Reports of Thinking in Multiple-Choice Critical Thinking Test Design

A methodology is described for using verbal reports of thinking to develop and validate multiple-choice tests of critical thinking. The procedure includes the following steps: devising normative models of thinking for each item; developing a method for interviewing subjects; collecting verbal reports of thinking from samples of subjects; using the normative models to rate the quality of the thinking portrayed in the reports; comparing the quality of thinking to answer choices and accumulating those comparisons across subjects for each item; identifying and modifying suspect items. The methodology is offered as a way to gather evidence on the truth of claims that currently available multiple-choice critical thinking tests do not measure critical thinking (McPeck, 1981; Petrie, 1986; Whimbey, 1985), and as a way to develop new critical thinking tests.

On présente ici une méthodologie d'utilisation des rapports verbaux sur la pensée pour développer et valider des tests à choix multiples de pensée critique. Le procédé comprend les étapes suivantes: conception de modèles normatifs de pensée pour chaque item; développement d'une méthode d'entrevue avec les sujets; collecte de rapports verbaux sur la pensée utilisant des échantillons de sujets; utilisation des modèles normatifs pour évaluer la qualité de la pensée telle qu'elle ressort des rapports; comparaison de la qualité de la pensée aux choix de réponses, et accumulation de ces comparaisons pour tous les sujets pour chaque item; identification et modification des items suspects. La méthodologie est offerte comme un moyen d'obtenir de l'évidence sur la valeur des idées selon lesquelles les tests à choix multiples actuels ne mesurent pas la pensée critique (McPeck, 1981; Petrie, 1986; Whimbey, 1985), et comme un moyen de développer de nouveaux tests de pensée critique.

During the last decade interest has grown in teaching and assessing critical thinking (Ennis & Norris, 1990). However, few assessment instruments are available. Furthermore, most of those available are multiple choice, and many advocates of critical thinking are dissatisfied in principle with multiple-choice testing formats (e.g., McPeck, 1981; Petrie, 1986; Whimbey, 1985). Such advocates insist that critical thinking tests must provide direct evidence on the thinking processes of examinees who take the tests, and multiple-choice critical thinking tests typically do not give such evidence. However, multiple-choice tests are popular because of their efficiency (Ennis & Norris, 1990) and are suited particularly well to some assessment tasks (Norris, 1989). Therefore, if multiple-choice testing of critical thinking could be justified (and seen as justified, because one cannot ignore perceptions) through design procedures that provided more direct evidence on the thinking processes that the tests

Stephen Norris is a professor of educational research and philosophy of education. In addition to critical thinking testing, his interests include the theory of reading comprehension, and philosophy of science and science education.

elicit, then both the critics and some important assessment goals might be satisfied.

This article describes how verbal reports of examinees' thinking on draft items of multiple-choice critical thinking tests can be used systematically to provide such direct validity evidence. Direct evidence on thinking processes gathered at the test design stage can be used to support more strongly inferences about examinees' thinking processes made at the use stage than design procedures that do not gather such direct evidence. Because direct evidence on thinking processes can play such an important role in test interpretation, testing theorists have endorsed verbal reports of thinking for amplifying the meaning of the constructs a test is measuring (Cronbach, 1971), for representing those constructs by the mental processes that underlie performance (Embretson, 1983; Embretson, Schneider, & Roth, 1986; Haney & Scott, 1987), for directly analyzing the processes underlying item performance (Messick, 1989), and for specifying the intellectual processes used to perform test tasks (Anastasi, 1988). Such endorsements support the premise of this article, namely, that the use of verbal reports of thinking can help meet the criticism that there is no direct evidence that multiple-choice critical thinking tests in fact measure critical thinking.

Despite the theoretical support, there are very few reported uses of verbal reports of thinking in test development and validation. When verbal reports have been used, the information in them rarely has been analyzed and compiled systematically. Rather, as will be shown in the section on past use of verbal reports of thinking in test design, the information usually has been used casually to support intuitive judgments about whether items mislead, contain unwanted clues, use unknown vocabulary, and the like. These are important considerations in test design, but it would be useful to see how verbal reports of thinking might be used systematically to help make construct validity judgments about multiple-choice critical thinking tests.

This article reports how verbal reports of thinking were used to develop and validate a multiple-choice critical thinking test of judging the credibility of observations. With suitable adjustments, the approach could be used to develop and to examine the validity of multiple-choice tests of other aspects of critical thinking.

Background

The Role of Observation in Critical Thinking

For the purposes of this study, Ennis's (1987) definition of critical thinking was adopted: "Critical thinking is reasonable reflective thinking that is focused on deciding what to believe or do" (p. 10). According to Ennis, thinking critically starts from a basis of information (information from others, from observation, and from conclusions previously drawn), and, through the process of inference, moves from that basis to decisions about belief and action. Observations are central to critical thinking on this model, because they are part of the basis on which decisions are based.

Motivation

Most definitions of critical thinking, including Ennis's, focus more on the process of thinking than its products. Paul (1985) argues that for learning to be

rational "the process that leads to belief is more important than belief itself" (p. 5) It is a result of this focus on process that many critical thinking theorists doubt that traditional multiple-choice tests can measure critical thinking, since such tests reveal products of thinking, not thinking itself.

This doubt is not pedantic. Criticisms of existing multiple-choice critical thinking tests have shown that keyed answers can be reached by thinking uncritically and unkeyed answers by thinking critically (Ennis & Norris, 1990; Norris, 1991). Such criticisms are usually hypothetical and not supported by empirical evidence on how examinees choose their answers. The point of the criticisms, however, is that there is no direct evidence one way or the other. Given the lack of evidence, there is a perceived unacceptable risk that some examinees are rewarded unfairly and others penalized unfairly. Such perceptions lead to a skepticism about the suitability of multiple-choice tests for measuring critical thinking.

A similar conclusion about multiple-choice critical thinking tests follows from research on thinking in complex domains and in ill-structured problems. Such thinking can follow multiple paths to multiple solutions (Resnick, 1987; Spiro, Vispoel, Schmitz, Samarapungavan, & Boerger, 1987), and often there is no expert consensus on either a correct path or a correct solution. Thus one might wonder how multiple-choice tests, which typically demand one solution, can test for critical thinking in complex domains.

Advocates of critical thinking are concerned, therefore, that multiple-choice testing cannot serve well the goal of critical thinking evaluation. On the other hand, multiple-choice tests seem an appropriate format for testing certain aspects of critical thinking, such as using the many principles for appraising the credibility of information, and there is considerable interest in students' ability to appraise information as opposed to merely locating it (e.g., National Assessment of Educational Progress, 1985). So it is important to know whether multiple-choice critical thinking tests can be designed to provide more direct evidence on the thinking that underlies performance. Such direct evidence on construct validity would place these tests on a sounder theoretical and empirical footing. Eliciting verbal reports of thinking on trial test items as described subsequently is one plausible way to collect the direct evidence desired.

Despite the references made previously to direct evidence, construct validity must be inferred (Cronbach, 1971). However, inferences about thinking processes based on item statistics and content analysis are *more* indirect and hence more untrustworthy than is desirable (Connolly & Wantman, 1964). Schuman (1966), for instance, has demonstrated that item performance can show excellent variation, correlate well with other item performances, and be related to other variables in relevant ways, even though the majority of examinees do not understand the item. As a result, some testing specialists recommend seeking more direct evidence on examinees' thinking than that contained in item statistics, such as evidence from examinees' verbal reports of their thinking as they work through test items (Anastasi, 1988; Haney & Scott, 1987) and from an "analysis of reasons" (Messick, 1989). Evidence of this sort can support inferences rarely supportable by item analysis data alone.

Eliciting verbal reports of examinees' thinking is an attempt to gain information on the mental processes leading to their performance on the test, that is, to learn the information, mental strategies, and reasoning principles they use in choosing answers. The aim is to look beneath the surface of answer choices in order to gain more direct insight into the reasons for the choices, so that more trustworthy inferences can be made about the causes of performance and, derivatively, about construct validity. In the specific context of testing for critical thinking, the aim would be to use information in examinees' verbal reports of thinking to judge whether it is differences in critical thinking, and not something else, that account for differences in scores. In order for the procedure to work, a way of eliciting verbal reports and of using the information contained in them needs to be designed and justified.

Past Use of Verbal Reports of Thinking in Test Design

Verbal reports of thinking have not been used often in test design. When such studies have been reported, direction on how the information in verbal reports might be used to quantify the quality of examinees' thinking, on how to make comparisons between the quality of thinking as displayed in examinees' verbal reports and their choices of answers, and on how to accumulate and coordinate these comparisons across several examinees to help judge the validity of tests has been inadequate.

Bloom and Broder (1950) were among the earliest users of verbal reports of thinking in test development. They asked a sample of college students to think aloud while they worked through questions on a problem solving test. The verbal reports yielded "relatively consistent and meaningful data from the majority of subjects" (p. 90) and helped to support inferences from the students' problem solutions to the thinking that led to the solutions. However, attempts were eschewed explicitly to formalize the procedures for gathering and using the verbal report information.

Kropp (1956) used verbal reports of thinking to examine the relationship between choosing correct answers on 10 multiple-choice items and the reasoning used to make the choices. The items were designed to measure understanding of text with physical science content and were judged to involve more complex thinking than simply recalling information or locating it in the text. The verbal reports showed that a great variety of reasoning processes led to choosing correct answers. Because some of these processes were not justifiable, such as outright guessing, Kropp concluded that it is hazardous to infer the quality of reasoning processes on the basis of answers chosen. However, no procedure was described for determining the *tendency* for unjustified reasoning to lead to correct answers across the sample of examinees for each item.

McGuire (1963) used verbal reports of thinking to construct medical school certification examinations that test for critical thinking and problem solving and not merely factual recall. Experts in the field were asked to classify test questions according to a modified Bloom's taxonomy. Medical students were then asked to think aloud as they worked through the items. McGuire reported that comparing the experts' classifications to the students' verbal reports helped her to align better the objectives of instruction with their assessment. However, it is unclear exactly how the verbal reports were ex-

amined and how the information in them was used to develop the test. As in the previously described studies, no specific guidance is given on how verbal reports of thinking can be used in test development.

Connolly and Wantman (1964) reported a study intended to improve that of Bloom and Broder (1950). Nine subjects were asked to report all the thoughts that might cross their minds as they worked on a set of 25 verbal analogy items. The adequacy of the subjects' thinking was compared with a model of ideal thinking on the items. The model was general rather than item specific and listed 67 behaviors representing quality of thinking (e.g., "Justifies choice with incorrect logic," and "Reads stem and immediately states keyed response"). The attempt to use a normative model of reasoning to judge verbal reports is significant for testing critical thinking, because critical thinking theory is a normative theory, that is, not merely a theory of thinking, but a theory of how to think *well*, namely, in accord with a set of standards that define how to think well (Ennis, 1981). However, Connolly and Wantman (1964) provided insufficient detail of what their ideal model is, where it originated, and how it was used. They concluded that their technique was useful for pretesting items, but did not mention a role for the technique in test validation. They suggested that the method might be used to determine whether some students are penalized for having too much knowledge. This suggestion is particularly relevant to multiple-choice critical thinking testing because of the allegations that it is knowledge recall and not critical thinking that these tests examine.

A study by Schuman (1966) provides the most detailed procedure for using the information in verbal reports of examinees' thinking. One thousand factory workers and cultivators in East Pakistan completed an attitude survey and then reported on the reasons for their responses to 10 randomly selected items. Their reasons were rated on a scale from 1, given to a reason that was clear and led to an accurate prediction of the chosen response, to 5, given to a reason that was unclear and could not be used to predict the answer chosen. Average scores were calculated for each item across all individuals and were used to indicate the subjects' understanding of that item: the lower the score the better the understanding. Schuman's procedure for quantifying and using the information in examinees' verbal reports marks an important difference in approach from previous studies, but the approach is not without problems. For instance, it was possible for examinees' reasoning to express clearly a *mis*understanding of an item, yet reliably lead to predictions of their responses. Their reasoning would be assigned 1, and the item would receive endorsement from a low average score, even though subjects did not understand it.

In sum, the research on using verbal reports of thinking for developing and validating tests is not adequate. Few systematic attempts have been made to use the information contained in verbal reports to quantify the quality of examinees' thinking and to accumulate the ratings of quality for each item across individuals. Furthermore, little emphasis has been placed on developing normative models against which reports of thinking might be compared. Without such models, the choice of keyed answers cannot be justified except by some atheoretical consensus procedure. Verbal reporting techniques have

been used rather casually in designing multiple-choice tests, with the possible exception of Schuman's study.

A Method for Eliciting and Using Verbal Reports of Thinking

To use systematically verbal reports of thinking in the development and validation of multiple-choice critical thinking tests, I am proposing a nine-step procedure as follows:

1. developing a set of critical thinking principles that items would serve to test;
2. devising initial trial items to test the principles;
3. developing normative models of thinking on each item;
4. developing an interviewing methodology for eliciting trustworthy reports of thinking;
5. collecting verbal reports of thinking on items;
6. scoring verbal reports using the normative models of thinking as a standard, and scoring answers chosen according to the answer key;
7. comparing thinking on items to performance on them and systematically accumulating these comparisons across examinees for each item;
8. modifying suspected items;
9. retrying the test in accord with steps 5 to 8.

Following is a demonstration of each of these steps in the development of a multiple-choice credibility judgment test.

Developing Principles of Observation Appraisal

In order to assess critical thinking and to compare the quality of examinees' verbal reports of thinking with their choice of answers, we need some justified standard for rating quality of thinking, that is, a set of norms against which thinking can be judged. The particular test being used as an illustration here, the Test on Appraising Observations (Norris & King, 1983), was designed to assess ability to appraise reports of observations. Therefore, to judge the quality of examinees' thinking on this test, a comprehensive and defensible set of principles for appraising observations is needed. Several sources provide such principles. For example, principles based on observers' conflict of interest, on being asked leading questions, and on the skill of the observer find support from longstanding judicial practice (Brooks, 1983). The principle that being asked leading questions lowers the credibility of answers is also supported by psychological research on eyewitness testimony (Lindsay, Wells, & Rumpel, 1981; Loftus, 1979; Wells, Ferguson, & Lindsay, 1981; Yarmey, 1979). Several principles are supported by commonsense psychology. For example, it makes sense to think that an observation is more believable when made by a more alert observer.

Ennis (1981) has compiled a comprehensive set of principles from a variety of such sources and Norris (1984) and Norris and King (1984) have extended Ennis's set. A sample of such principles is contained in Table 1. The principles do not provide either necessary or sufficient conditions for observation statements to be credible. They suggest conditions to consider. To the extent that these conditions are satisfied, an observation is worthy of belief, and to the extent they are not satisfied, it is worthy of doubt. In any given case, principles

Table 1
Illustrative Principles for Appraising Observations Selected from a Comprehensive Set by Norris and King (1984)

I. Observation statements tend to be more credible than inferences based upon them.
II. An observation statement tends to be credible to the extent that the *observer:*
 1. is alert to the situation and gives his or her statement careful consideration;
 2. has no conflict of interest;
 3. is skilled at observing the sort of thing observed.
III. An observation statement tends to be credible to the extent that the *observation conditions:*
 1. provide a satisfactory medium of observation;
 2. provide sufficient time for observation;
 3. include adequate instrumentation, if an instrument is used.
IV. An observation statement tends to be credible to the extent that the *observation statement:*
 1. is corroborated;
 2. is made close to the time of observing;
 3. is not given in response to a leading question.

may compete, and there are no strict procedures for arbitrating such difficult cases. Good judgment and experience are needed.

Devising Initial Trial Items

Once a set of principles is compiled and defended, items can be constructed to test knowledge of them. The construction requires that situations calling for judgments using the principles be devised, and that the conclusions implied by the principles in these situations be determined. These conclusions then serve as a key, and examinees' knowledge of the principles can be inferred from whether their conclusions match those in the key.

Based on this general scheme, initial trial items were written. They were in multiple-choice format and each was written to test ability to use one principle of observation appraisal. Fifty items were cast in the context of two stories: Part A, the story of a traffic accident, and Part B, the story of a river exploration. The story context was used in order to provide examinees sufficient context for deciding which principle of observation appraisal should be applied.

Each item presented two conflicting reports of observations made by characters in the stories and required examinees to decide which, if either, of the reports was more credible. Comparative judgments were requested because, without the wealth of information available in real situations, usually it is not possible to make absolute judgments about the credibility of single observation statements (Ennis, 1988; Norris & Ennis, 1989). Thus for each item the three possible answer choices were: the first statement is more credible, the second statement is more credible, and the statements are equally credible. Multiple-choice items often have more than three alternatives to reduce the likelihood of guessing correctly, but the above three are the only logically obvious ones given the task.

Developing Normative Models of Thinking for Items

The principles of observation appraisal were used to develop a normative model of thinking on each item, that is, a model that instantiates one of the principles of observation appraisal and uses the principle to support a judgment of credibility. As an example, consider Item 3 from Part A, the traffic accident story:

> A policewoman has been asking Mr. Wang and Ms. Vernon questions. She asks Mr. Wang, who was one of the people involved in the accident, whether he had used his signal.
> Mr. Wang answers, "Yes, *I did use my signal.*"
> Ms. Vernon had been driving a car which was not involved in the accident. She tells the officer, *"Mr. Wang did not use his signal.* But this didn't cause the accident."

Examinees were to choose which, if either, of the emphasized statements was more credible. To reason through this question correctly, an examinee first needs to be able to derive from the text the relevant information about Wang's and Vernon's involvement. The text is simple enough that most high school students should have no difficulty with this aspect of the task. Second, an examinee must retrieve from background knowledge the relevant facts that not using a turn signal can cause an accident and that being held responsible for an accident can cause one trouble with the law. Again, high school students would have such common knowledge. Finally, an examinee has to infer that the possibility that one's own testimony can place one in trouble with the law is an accuracy-reducing factor, and thus that Wang is less credible than Vernon.

The normative model of thinking developed for this item is as follows:

> Mr. Wang was involved in the accident, but Ms. Vernon was not involved. Mr. Wang is less credible because his involvement would give him reason to say he used his signal even if he did not. Wang is in a conflict of interest. People in a conflict of interest, that is, people who have something to gain by events being cast as they described them, tend to be less credible than those who are not in such a situation.

This normative model instantiates the principle of conflict of interest by showing how Wang is in such a conflict, and then uses the principle to support the judgment that Wang is less credible than Vernon.

The reasoning required for Item 3 generalizes to the following three steps that are required for sound reasoning on all the test items:

1. cite the relevant facts in the text that can be used to compare the credibility of the emphasized statements;
2. use these facts together with relevant background knowledge to make a comparative evaluation of the credibility of the statements;
3. show how the evaluation is based on an appropriate general principle that subsumes the relevant facts in the text under a general accuracy-reducing or accuracy-increasing factor.

Each item on the test was analyzed according to this scheme, normative models of reasoning for each item developed, and the models used in a

manner described subsequently to rate the quality of examinees' verbal reports of thinking.

Developing an Interviewing Methodology

An interviewing methodology was designed to elicit examinees' reasoning while they worked through items on the observation test. In accord with widely advocated practices for collecting verbal reports as data (Afflerbach & Johnston, 1984; Ericsson & Simon, 1980, 1984; Larkin & Rainard, 1984; Loftus, 1979; Norris, 1991) the interviews were as nonleading as possible. The initial directive to examinees was: "As you do each question tell me all you can about what you are thinking while you are picking your answer." At this stage, interrupting an examinee's verbal reporting was permitted only to clarify the ambiguous referent of a pronoun or to point to obvious reading errors.

It is important to obtain records of reasoning that are as complete as possible. To fulfill this aim it is necessary sometimes to probe beyond the initial instruction to think aloud, while being careful not to rush examinees by prematurely cutting off their reasoning (Larkin & Rainard, 1984) or to endorse or criticize particular reasoning attempts. Even in such follow-up stages, it is desirable for probing to be as nonleading as possible, by merely echoing examinees' reported thoughts or by asking them to explain a little more what they have said.

Some researchers who use this method think that some training is needed for subjects to give useful verbal reports of their thinking. To this end, such techniques as using demonstration tapes (Bridge & Winograd, 1982), practice verbal reporting sessions (Kavale & Schreiner, 1979), practice examples (Larkin & Rainard, 1984), and prestudy introspection (Johnston & Afflerbach, 1983) have been tried. However, it is not known what effect, if any, such practice has on the verbal reports that subjects subsequently give (Afflerbach & Johnson, 1984). For this reason, and because of the amount of time that we asked subjects to devote to the actual data gathering phase, training sessions were not used in the research reported here.

To meet the above goals and constraints, the three-stage interview methodology given in Appendix A was adopted. In Stage One the interviewer informs examinees of the general purpose of the interviews and of their role in them, especially of the fact that they will be asked some questions about their thinking to be answered verbally while taking a test.

In Stage Two the interviewer asks examinees to tell all they can about what they are thinking while choosing their answers. Interruptions to examinees' reports are permitted only to probe for ambiguous references and for obvious reading errors (see II.2.1 and II.2.2). If examinees ask for additional information or for reasons, the interviewer must respond as in II.3.1, "You can only go by what is written," or II.3.2, "You can decide only according to what is said and what you know." If examinees seek feedback on their progress it must not be given (II.4.3).

Stage Three begins after examinees have chosen their response and have finished reporting on their thinking in response to the directive in II.1. The interviewing procedure for this stage consists of a set of condition-action pairs that guide the actions of the interviewer in accord with the satisfaction of conditions defined by a combination of examinees' answer choices and the

content of their verbal reports. Each action is designed to match a particular combination by being as nonleading as possible and still having a chance to elicit useful information. For any given item and examinee response, only one numbered step applies. The strategies listed under III.4 guide the interviewer when it is not clear how to proceed given an examinee's report.

Stage Three proceeds as follows. Starting with III.1 and proceeding sequentially, the interviewer checks whether the condition (described by the phrase following the *if*) is satisfied. Sometimes the condition is complex, that is, consists of subconditions, so all subconditions must be met for the condition to be satisfied. If a condition is not satisfied, the interviewer proceeds to the next numbered step. If a condition is satisfied, the interviewer carries out the indicated action (described by the phrase following the *then*). When the signal *then proceed to the next item* is reached the interview reverts to Stage Two.

Collecting Verbal Reports of Thinking

Our experience has shown that a multiple-choice test takes about four times as long in a verbal reporting format as in a paper-and-pencil format. Therefore, it is too tiresome for students to be interviewed on a whole test designed for a regular class period when taken in paper-and-pencil format. However, if there are story lines in a test and items are linked thematically, as is the case with the Test on Appraising Observations, or if the items are in a logical sequence for another reason, a student could not be interviewed on a later section without having completed previous sections. A procedure for interlacing interviewing and paper-and-pencil formats is needed.

Subjects whose verbal reports of thinking were to be collected were divided randomly into halves; one half was assigned to take Part A of the test and the other half to take Part B. Each part of the test was divided into two sections representing logical breaks in the story line: Part A, Section I (Items 1-15) and Section II (Items 16-28); Part B, Section I (Items 29-37) and Section II (Items 38-50). Half of the subjects assigned to a part were randomly chosen to be interviewed on Section I of that part and to complete Section II in paper-and-pencil format; the other half was to write Section I and be interviewed on Section II.

An interviewer dealt with two students at a time working on the same part, either A or B, of the test. First, both students were told how the testing would proceed. The purpose of the study had already been explained in meetings with students in their classes. The student who was to write Section I of the part was taken to a place to work alone, instructed to stop when finished the last item of Section I (either Item 15 or 37), not to read on, and to wait to be called by the interviewer. The other student was interviewed on Section I according to the methodology described in Appendix A. When the interview was over, the students exchanged places. The student who wrote Section I was interviewed on Section II and the student who was interviewed on Section I wrote Section II. Students marked their answers to all questions on a standardized answer sheet, and students' verbal reports were tape-recorded.

The results of two such interviewing stages, for two sequential versions of the Test on Appraising Observations (Versions B and C), are described and compared in this article. For version B, 52 grades 10 and 11 students from seven small rural schools took part. For Version C, 45 students from four

Table 2
Numbers of Examinees Giving Verbal Reports on Each Test Section for Versions B and C of the Test on Appraising Observations

Version	Part A		Part B	
	Sect. I	Sect. II	Sect. I	Sect. II
B	13	14	13	12
C	13	12	10	10

different rural schools were involved. Students were selected randomly from class lists and given the opportunity to participate. None refused. Verbal reports for each test section were collected on the number of students shown in Table 2.

Scoring Verbal Reports of Thinking and Answer Choices

Performance scores were assigned to answer choices for each item according to whether they agreed (1) or disagreed (0) with a key based on the set of principles illustrated in Table 1.

Verbal reports were transcribed verbatim and analyzed. Each student was assigned a *thinking score* for each item according to the rating scale in Table 3. The rating scale is based on the normative models of thinking and assigns scores according to how closely students' verbal reports approximate the models. Students were not privy to the rating scale. Thus, the amount of detail and generality provided in students' verbal reports was entirely up to them. The aim was to determine what students would say when left to their own resources, since this best simulated the paper-and-pencil text administration. Table 4 shows how the rating scale would work for Item 3.

For illustration, here are the verbatim transcriptions of the verbal reports of two students working on Item 3:

Student A ... ah ... ah, Mr. Wang, like he probably didn't, like you know, it was just, he probably thought he used his signal, but really didn't. And Miss Vernon, she was watching, so she'd be able to

Table 3
Rating Scale for Assigning Thinking Scores

Rating	Basis of Evaluation
1	The examinee indicates a criterion by which a correct comparison of the emphasized statements might be made, or the examinee uses a criterion in comparing the two statements but does not explicitly indicate what the criterion is.
2	The examinee indicates a criterion and correctly uses it to compare the emphasized statements.
3	The examinee indicates a criterion, correctly uses it to compare the emphasized statements, and justifies the comparison on the basis of a general principle.
0	The examinee does none of the above, does not think critically in some other way, or does not respond.

Table 4
Assigning Thinking Scores for Item 3

Rating	*Basis of Evaluation*
1	The examinee points out that Mr. Wang was involved in the accident.
2	The examinee points out that Mr. Wang was involved in the accident and compares Mr. Wang's involvement to Ms. Vernon's being a bystander.
3	The examinee points out that Mr. Wang was involved in the accident, compares this with Ms. Vernon's noninvolvement, and shows that these facts put Wang in a position of people who stand to profit or lose depending upon what they say.
0	The examinee does none of the above, does not think critically in some other way, or does not respond.

tell from back if he was using it or not. Right? Being the case, so, I'd tend to believe Vernon.

Student B I would say that he did use his signal because anybody who's in the car ... coming up to an intersection or anything ... he, he usually knows what he's doing. So I'd be more inclined to believe the first.

Student A chose the keyed response and Student B an unkeyed one. Neither student thought critically on the item, however, and both were assigned zero for thinking scores. The criteria used by both students for comparing Wang and Vernon were unsound. Student A's reasoning that Wang just "thought he used his signal," but Vernon would "be able to tell from back" because "she is watching" is arbitrary. There is no information in the item to justify the claims about either Wang or Vernon. It is used to rationalize a choice of answer rather than to justify it. Given the information in the item, it is just as reasonable to say that Wang would "be able to tell from inside that he used his signal" and that Vernon just "thought he did not use his signal." Student B's reasoning fails to distinguish between Wang's *knowing* what he had done and his *reporting accurately* what he had done, and to allow that someone in another car can know what another driver is doing. The students' performance scores and thinking scores for Item 3 are as in Table 5.

For each student, there was a performance score for each item taken and, for each item done in the interview, a thinking score also. Performance and thinking scores were conceptually and empirically independent: there is little difficulty in seeing that although Student A chose the keyed answer, the student's justification was poor; across the two students, performance and

Table 5
Two Students' Performance and Thinking Scores for Item 3

Student	*Performance Score*	*Thinking Score*
A	1	0
B	0	0

thinking scores are uncorrelated. Thinking scores were assigned without reference to answer chosen. When appropriate, students were given credit for using correctly criteria other than those assumed when designing the items. That is, the ideal model of thinking for each item was used as a guide, but it was possible for reasoning to be rated highly according to general standards of critical thinking even though it was outside the bounds of the specific normative model on which this test is based. For instance, an examinee might reject a factual presupposition built into an item and reason perfectly acceptably from an alternative presupposition to other than the keyed response. In such a situation, the thinking score would be high, but the performance score would be zero.

Comparing Performance and Thinking Scores and Judging Items

A central feature of the test development and validation methodology is the role that the relationship between performance and thinking scores plays in the evolution of a test. The aim is to adjust items, or to write new ones, so that the relationship between quality of thinking and quality of performance is strong for each item. The relationship was calculated using point biserial correlation coefficients between performance scores and thinking scores for each item across subjects sampled. Consider the following two sets of ordered pairs of thinking and performance scores for Items 18 and 38 in Version B of the test:

{18 | (2,1)(0,0)(0,0)(2,1)(2,1)(2,1)(2,1)(0,0)(2,1)(2,1)(1,1)(2,1)(2,1)(2,1)};

{38 | (0,1)(0,0)(0,1)(0,0)(0,0)(0,0)(1,0)(0,0)(0,1)(0,0)(1,0)(1,1)}.

For Item 18 all the thinking/performance combinations support the validity of the item. That is, for each subject, choosing the keyed answer is associated with thinking well and choosing an unkeyed answer is associated with thinking poorly. This relationship is reflected in a high thinking/performance point biserial correlation of 0.96. Comparing Item 38, for which there are five mismatches between thinking and performance scores (either combinations [1,0] or [0,1]), the computed point biserial correlation is 0.33. This evidence suggests that Item 18 is suitable as it stands, but that Item 38 needs modification. An examination of traditional item statistics, difficulty levels, and item/test biserial correlations also suggest Item 38 needs modification. For Items 18 and 38, respectively, we have: difficulties of 0.56 and 0.46; item/test biserials of 0.45 and 0.21.

However, this correspondence between traditional indicators of item validity and the thinking/performance indicator recommended here did not obtain for all items. For instance, Item 5 on Version B had a very high thinking/performance point biserial of 0.89, but an item/test biserial of 0.19. Placing more emphasis on the thinking/performance relationship for items is an attempt to meet the challenge from theorists of critical testing that traditional measures of internal consistency (KR-20 reliability and item/test biserials) may not suit critical thinking testing, because critical thinking is a heterogeneous concept. The approach here was to treat the item/test biserial as a dichotomous indicator—only if the item/test biserial for an item was negative was this indicator used to call an item into question.

Modifying Suspected Items

As described in the previous section, item level statistics, such as item difficulty levels, item/test biserial correlations, and thinking/performance correlations were used to identify problematic items in experimental versions of the test. Most weight was given to the thinking/performance correlation. There were 19 items on Version B with thinking/performance correlations <0.5 (rounded to one decimal place). This figure was taken as the lower bound of acceptability, and reasons were sought for these low indexes by examining the transcribed interviews and asking the following questions:

1. Do examinees understand the task in the intended way; or are the instructions unclear, is the item systematically misleading, or is the assumed background knowledge outside their knowledge store?
2. Are examinees thinking critically to arrive at other than the keyed response and, if so, is it the keyed response or is it the item that should be changed?
3. Are examinees thinking uncritically yet arriving at the keyed answer and, if so, are there clues in the text that can be eliminated?

For Version B, all but four of the 19 suspect items, items 12, 17, 22, and 23, were altered on the basis of information contained in the verbal reports. Alterations included single word changes, the addition and deletion of information, and changes to the directions. Items 12, 17, 22, and 23 were not altered because there was no clear basis for alteration contained in the verbal reports and because it was expected that changes to neighboring and thematically related items would improve them also.

An example of changes made to the directions serves to illustrate how the thinking/performance correlations and the verbal reports were used together. Recall that Part A of the test was set in the context of a traffic accident. The Version B directions listed the names of all characters and what they were doing at the time of the accident. With several characters in the story in different roles, this list was provided to help examinees keep them straight. But the verbal reports showed that many examinees looked for direct answers to questions in this information. In one item, for example, characters gave conflicting reports about how many cars were at the intersection when the accident occurred. One character was more alert to the traffic and hence a more credible source of information about the number of cars at the intersection. Coincidentally, that character reported seeing the same number of cars *at the intersection* (3) as the directions mentioned were *involved in the accident*. The verbal reports showed that several examinees did not consider the characters' alertness, but merely counted the number of cars mentioned in the directions and chose the keyed answer as a result. These students thought uncritically, because they equated the number of cars at the intersection with the number involved in the accident, but chose the keyed response. It is the possibility of just this sort of occurrence with multiple-choice tests that advocates of critical thinking find most disturbing.

Retrying the Modified Test

When the 19 suspected items were examined and 15 of them revised, the new version (Version C) was tried by collecting verbal reports of thinking on it, calculating thinking/performance correlations, and so on. Figure 1 displays a comparison of the thinking/performance correlations for Versions B and C.

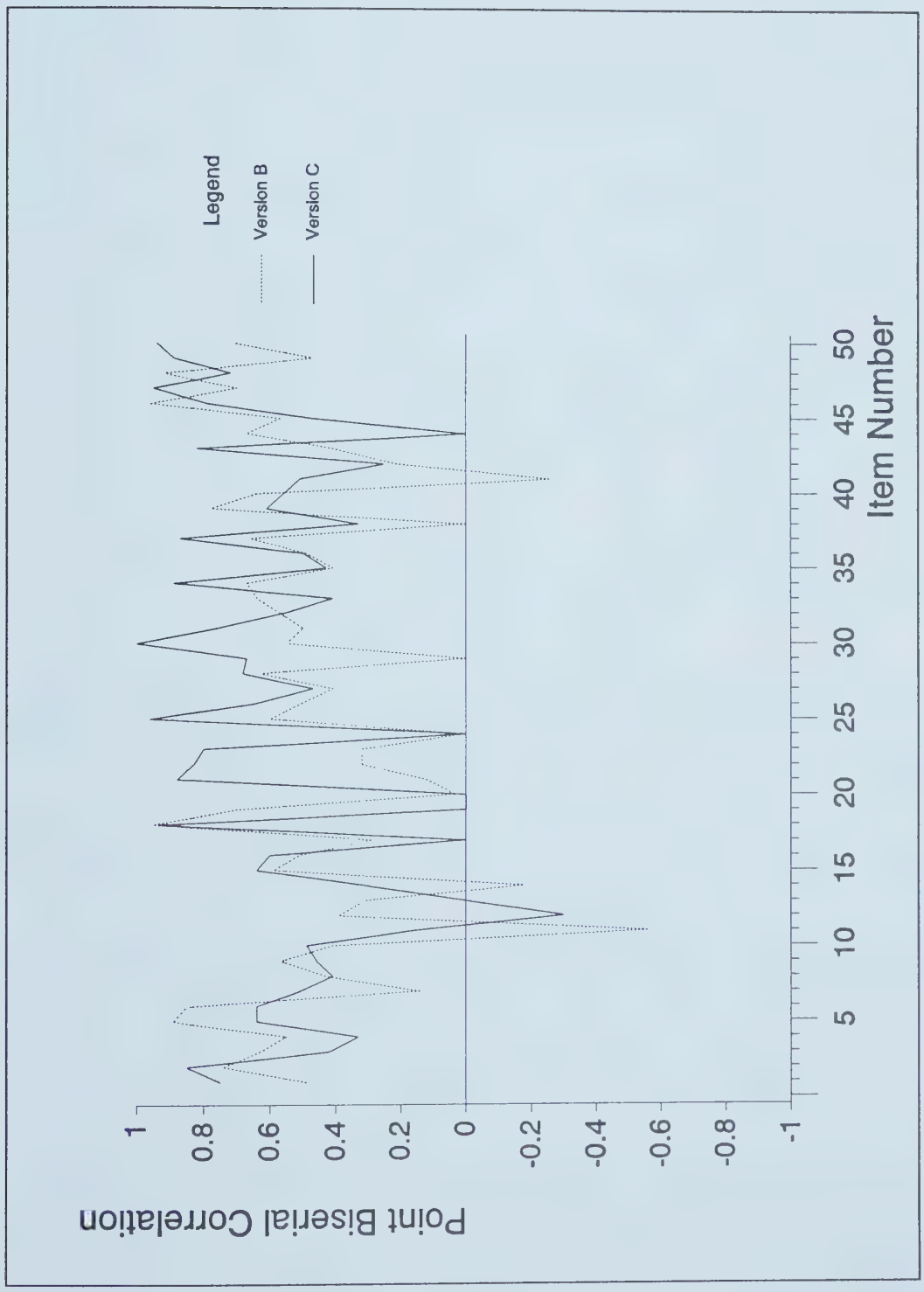


Figure 1. Thinking/performance correlations by item for Versions B and C.

There was a general increase in thinking/performance correlations for Version C: 29 items in Version C had larger thinking/performance correlations than Version B; 20 items had smaller correlations; and one correlation was equal for the two versions. The most consistent increases were made in the range of Items 21 to 37. Most telling, there was still one negative correlation, for Item 12, and six zero correlations. For three of the zero correlations, all examinees answered the item correctly and had thinking scores greater than zero. The items were too easy for this sample. In the other three, all examinees received zero for thinking and up to four students chose the keyed answers. These items were too difficult, but, many would argue, should remain on the test because they represent critical thinking principles that subjects should know.

Other items on Version C clearly needed further revision and trial. The interviews indicated for items 3, 4, and 8 a common error in students' thinking that was largely responsible for the low thinking/performance correlations. In each item, several students made their choice solely on the basis of who was the driver of a car on the grounds that a driver would be more alert to the traffic than a pedestrian or passenger. For example, in Item 3 on Version C, Mr. Wang was a driver and Ms. Vernon "watched the accident happen." Several students chose Mr. Wang's statement as more credibile because he was a driver and thus more alert to road conditions. These students were assigned thinking scores of 1 for recognizing this important factor, even though they neglected the more important fact that Wang was in a conflict of interest. Thus they chose an unkeyed response but thought fairly well, contributing to a lowered thinking/performance correlation. The obvious change was to make both Wang and Vernon drivers, as is the case in the version of Item 3 cited previously as an example. Items so modified require retrial following the procedures followed in modifying Version B.

Discussion and Conclusions

Verbal reports of thinking on trial items of multiple-choice critical thinking tests seem useful and important from at least three perspectives. First, existing multiple-choice critical thinking tests are often criticized because they give no direct evidence of the process of thinking that leads to examinees' choices of answers. The criticism is justified and important, because evidence on the thinking processes used to answer items has a direct bearing on construct validity. If examinees choose keyed answers but they do not think critically, or choose unkeyed answers but think critically, then the test is not valid.

Second, multiple-choice critical thinking tests could serve several useful purposes if they were valid. They are useful when testing groups of students, when time for grading is at a premium, and when knowledge of several criteria and principles of critical thinking is being tested. That is, multiple-choice critical thinking tests can fill a particular set of roles, though they probably cannot serve all the needs of critical thinking assessment. For instance, multiple-choice tests probably cannot assess ability to think critically in complex situations where solutions are radically underdetermined by available information; they cannot test for the ability to orchestrate many critical thinking abilities; and they cannot test for critical thinking dispositions.

Third, direct evidence on the information, strategies, and principles that examinees use to answer items on existing multiple-choice critical thinking

items would be directly relevant to the construct validity of those tests. In the context of test design, direct evidence from examinees' verbal reports can be used to discard, modify, and retain items as done in this study. In the context of test use, the test would inherit the weight of this direct evidence gathered in the design context, even though in the use context no direct evidence would be gained on the causes of performance. In addition to being useful in test development, the verbal reporting procedure would be useful for evaluating the claims that currently available multiple-choice critical thinking tests do not measure critical thinking (Ennis & Norris, 1990; McPeck, 1981; Petrie, 1986; Whimbey, 1985). For instance, the procedure could be adapted for studying the validity of the Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 1980) and the Cornell Critical Thinking Test Levels X and Z (Ennis & Millman, 1985), two of the oldest and most widely used multiple-choice critical thinking tests.

A number of possible outcomes emerge from such studies. First is the possibility that the criticisms of these tests would be sustained. This would be unfortunate, because it would cast doubt on much of the research on the effectiveness of critical thinking instruction that used the tests as criterion measures. Another possible result is a clearer understanding of the range of applicability of existing critical thinking tests. Many multiple-choice critical thinking tests are advertised for use over wide age ranges. For example, the Cornell Critical Thinking Test Level X is said to be usable from grade 4 to grade 14 and the Watson-Glaser Critical Thinking Appraisal from grade 9 up. These ranges include individuals who differ widely in intellectual sophistication, empirical knowledge of the world, and test-taking experience, so it is quite plausible that the tests are not equally suitable across the entire ranges. A study of the verbal reports of thinking of examinees from different groups would provide direct evidence on the range of suitability of each test. A possible consequence is that the range for each test will need to be narrowed. Another possibility is that answer keys might be adjustable to match more suitably the critical thinking of examinees of different age levels.

The verbal reporting approach itself needs further study using different types of test content, different elicitation procedures, and different types of examinees. It is known, for instance, that different questioning procedures can affect differently the reports eyewitnesses give of events (Loftus, 1979). Can different approaches to eliciting verbal reports of thinking alter what examinees report? One study (Norris, 1990) shows that verbal reports of high school students' thinking are not easily altered by asking them more leading questions, and that performance scores while giving verbal reports are the same as performance scores in a paper-and-pencil test taking situation. This is promising. But would verbal reporting work as well with other groups and in other testing situations?

Indirect evidence for the value of the approach can be found by examining standard indicators of test validity and reliability. The Cornell Critical Thinking Test (Ennis & Millman, 1985) and the Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 1980) provide criteria for judging the validity of the Test on Appraising Observations and hence of indirectly judging the effectiveness of the verbal reporting technique. Those tests have 71 and 80

items respectively, whereas the Observations test has 50 items. This would tend to make the Observations test less reliable. Comparison of the mean KR-20 reliabilities for the three tests computed from the performance of four groups of senior high school students from Southern Ontario (N=100, 101, 106, 287) shows that the reliability of the Test on Appraising Observations is only slightly less than the Cornell (0.69 vs. 0.72). The Watson-Glaser reliability of 0.80 is about what would be predicted by increasing the number of items from 50 to 80 (see Lord & Novick, 1968, Fig. 5.10.1, p. 113). Thus reliability does not seem to have been sacrificed to validity in using verbal reports and focusing on item validity to develop the observations test.

It was also expected that the correlation of the Observations test with the Cornell test should be higher than with the Watson-Glaser test. The rationale for this expectation is that the Cornell test includes a section of 24 items on judging the credibility of reports of observations, but the Watson-Glaser has no items on judging credibility. Using the data on the same four groups of Southern Ontario students, the average correlation between the Observations and Cornell test was 0.48 ($p<.01$), while the average correlation between the Observations and the Watson-Glaser test was 0.31 ($p<.01$), in line with predictions. Also in line with predictions, for a group of 172 grades 10, 11, and 12 students with a wide range of academic ability attending a large metropolitan school in the Atlantic provinces, the correlation between the Observations and Cornell tests was 0.62 (Bonferroni adjusted $p<.05$), and the correlation between the Observations and Watson-Glaser tests was 0.35 (Bonferroni adjusted $p<.05$).

The verbal reporting methodology thus led to a test that meets several expectations. This is evidence for the value of the methodology. Therefore, eliciting verbal reports of examinees' thinking on multiple-choice critical thinking tests, comparing the information in the verbal reports to examinees' choices of answers, and using these comparisons systematically to weigh and balance the evidence for the quality of items seems a valuable approach for developing and validating such tests. In particular, it is a way to satisfy two competing desires of many proponents of critical thinking: the desire to use multiple-choice critical thinking tests because they provide a convenient way to gather information on students' critical thinking and the desire to have information on the process, not just the products, of that thought.

Acknowledgments

This research was supported by grants from the Social Sciences and Humanities Research Council of Canada (Grant Nos. 418-81-0781 and 410-89-0052).

I thank Linda M. Phillips and an anonymous reviewer for helpful comments.

References

Afflerbach, P., & Johnston, P. (1984). On the use of verbal reports in reading research. *Journal of Reading Behavior, 41*(4), 307-322.

Anastasi, A. (1988). *Psychological testing*. New York: Macmillan.

Bloom, B.S., & Broder, J.L. (1950). *Problem-solving processes of college students*. Chicago: University of Chicago Press.

Bridge, C., & Winograd, P. (1982). Readers' awareness of cohesive relationships during cloze comprehension. *Journal of Reading Behavior, 14*, 299-312.

Brooks, N. (1983). *Police guidelines: Pretrial eyewitness identification*. Ottawa: Law Reform Commission of Canada.

Connolly, J.A., & Wantman, M.J. (1964). An exploration of oral reasoning processes in responding to objective test items. *Journal of Educational Measurement, 1*, 59-64.
Cronbach, L.J. (1971). Test validation. In R.L.Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 443-507). Washington, DC: American Council on Education.
Embretson, S. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin, 93*, 179-197.
Embretson, S., Schneider, L.M., & Roth, D.L. (1986). Multiple processing strategies and the construct validity of verbal reasoning tasks. *Journal of Educational Measurement, 23*, 13-32.
Ennis, R.H. (1981). Rational thinking and educational practice. In J. Soltis (Ed.), *Philosophy and education . Eightieth yearbook of the National Society for the Study of Education, Part 1* (pp. 143-183). Chicago: NSSE.
Ennis, R.H. (1987). A taxonomy of critical thinking dispositions and abilities. In J.B. Baron & R. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9-26). New York: Freeman.
Ennis, R.H. (1988). Testing teachers' competence, including their critical thinking ability. In B. Arnstine & D.G. Arnstine (Eds.), *Philosophy of Education 1987* (pp. 413-420). Normal, IL: Philosophy of Education Society.
Ennis, R.H., & Millman, J. (1985). *Cornell critical thinking test, level X*. Pacific Grove, CA: Midwest Publications.
Ennis, R.H., & Norris, S.P. (1990). Critical thinking assessment: Status, issues, needs. In S. Legg & J. Algina (Eds.), *Cognitive assessment of language and math outcomes* (pp. 1-42). New York: Ablex.
Ericsson, K.A., & Simon, H.A. (1980). Verbal reports as data. *Psychological Review, 87*, 215-251.
Ericsson, K.A., & Simon, H.A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
Haney, W., & Scott, L. (1987). Talking with children about tests: An exploratory study of test item ambiguity. In R.O. Freedle & R.P. Duran (Eds.), *Cognitive and linguistic analyses of test performance* (pp. 298-368). Norwood, NJ: Ablex.
Johnston, P., & Afflerbach, P. (1983, December). *The construction of main idea in reading comprehension*. Paper presented at the annual meeting of the National Reading Conference, Austin, Texas.
Kavale, K., & Schreiner, R. (1979). The reading processes of above average and average readers: A comparison of the use of reading strategies in responding to standardized comprehension measures. *Reading Research Quarterly, 15*, 102-128.
Kropp, R.P. (1956). The relationship between process and correct item responses. *Journal of Educational Research, 49*, 385-388.
Larkin, J.H., & Rainard, B. (1984). A research methodology for studying how people think. *Journal of Research in Science Teaching, 21*(3), 235-254.
Lindsay, R.C.L., Wells, G.L., & Rumpel, C.M. (1981). Can people detect eyewitness-identification accuracy within and across situations? *Journal of Applied Psychology, 66*, 79-89.
Loftus, E.F. (1979). *Eyewitness testimony*. Cambridge, MA: Harvard University Press.
Lord, F.M., & Novick, M.R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
McGuire, C. (1963). Research in the process approach to the construction and analysis of medical examinations. *National Council on Measurement in Education Yearbook, 20*, 7-16.
McPeck, J. (1981). *Critical thinking and education*. New York: St. Martin's Press.
Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan.
National Assessment of Educational Progress. (1985). *The reading report card: Progress toward excellence in our schools: Trends in reading over four national assessments, 1971-1984*. Princeton, NJ: Educational Testing Service.
Norris, S.P. (1984). Trying eyewitness testimony, validating tests, and teaching critical thinking. In R.E. Roemer (Ed.), *Philosophy of Education 1983* (pp.179-189). Normal, IL: Philosophy of Education Society.
Norris, S.P. (1989). Can we test validly for critical thinking? *Educational Researcher, 18*(9), 21-26.
Norris, S.P. (1990). Effect of elicting verbal reports of thinking on critical thinking test performance. *Journal of Educational Measurement, 27*, 41-58.
Norris, S.P. (1991). Informal reasoning assessment: Using verbal reports of thinking to improve multiple-choice test validity. In J.F. Voss, D.N. Perkins, & J. Segal (Eds.), *Informal reasoning and education* (pp. 451-472). Hillsdale, NJ: Erlbaum.
Norris, S.P., & Ennis, R.H. (1989). *Evaluating critical thinking*. Pacific Grove, CA: Midwest.

Norris, S.P., & King, R. (1983). *Test on appraising observations*. St. John's, NF: Institute for Educational Research and Development, Memorial University of Newfoundland.
Norris, S.P., & King, R. (1984). Observation ability: Determining and extending its presence. *Informal Logic, 6*(3), 3-9.
Paul, R. (1985). Bloom's taxonomy and critical thinking instruction. *CT News, 3*(5), 3-6.
Petrie, H. (1986). Testing for critical thinking. In D. Nyberg (Ed.), *Philosophy of Education 1985* (pp. 3-19). Normal, IL: Philosophy of Education Society.
Resnick, L.B. (1987). *Education and learning to think*. Washington, DC: National Academy Press.
Schuman, H. (1966). The random probe: A technique for evaluating the validity of closed questions. *American Sociological Review, 31*, 218-222.
Spiro, R.J., Vispoel, W.L., Schmitz, J.G., Samarapungavan, A., & Boerger, A.E. (1987). Knowledge acquisition for application: Cognitive flexibility and transfer in complex content domains. In B.C. Britton & S. Glynn (Eds.), *Executive control processes* (pp. 177-199). Hillsdale, NJ: Erlbaum.
Watson, G., & Glaser, E.M. (1980). *Watson-Glaser critical thinking appraisal*. Cleveland, OH: Psychological Corporation.
Wells, G.L., Ferguson, T.J., & Lindsay, R.C.L. (1981). The tractability of eyewitness confidence and its implications for triers of fact. *Journal of Applied Psychology, 66*, 688-696.
Whimbey, A. (1985). You don't need a special "reasoning" test to implement and evaluate reasoning training. *Educational Leadership, 43*(2), 37-39.
Yarmey, A.D. (1979). *The psychology of eyewitness testimony*. New York: Free Press.

Appendix A
Interviewing Methodology for Eliciting Examinees' Reasoning on Multiple-Choice Tests of Critical Thinking (with Examples Based on the Test on Appraising Observations)

Stage One

I.1 Inform examinee of purpose of interview: to find out what he or she is thinking while choosing answers to questions on the test.

I.2 Inform examinee of his or her role: to respond as completely as possible to the questions asked on the test.

Stage Two

II.1 Interviewer says to examinee:
"As you do each question tell me all you can about what you are thinking while you are picking your answer."

II.2 Interviewer can interrupt examinee's narrative only to:

II.2.1 probe for ambiguous reference of demonstratives or third person pronouns by saying: "Could you tell me what you mean by ...?"
Example: When doing Item 3 the examinee says: "This driver here is more believable." Probe immediately: "Could you tell me what you mean by 'this driver here'?"

II.2.2 probe for obvious reading mistakes by saying:
"Did you read ...?" (Do not endorse answers.)
Example: When doing Item 3 the examinee reads: "Mr. Wang did use his signal" for "Mr. Wang did not use his signal." Probe immediately: "Did you read 'Mr. Wang *did* use his signal'?"

II.3 Interviewer can respond to examinee's questions only as follows:

II.3.1 If examinee probes for facts, say: "You can only go by what is written."
Example: When doing Item 3 the examinee asks: "Was Vernon close enough to see Wang's car?" Answer only: "You can only go by what is written."

II.3.2 If the examinee probes for reasons, say: "You can decide only according to what is said and what you know."

Example: When doing Item 3 the examinee asks: "How did Vernon know what caused the accident?" Answer only: "You can decide only according to what is said and what you know."

II.4 General cautions:

II.4.1 Do not begin to speak immediately after the examinee stops; give the examinee a few seconds to continue.

II.4.2 Do not cut off examinee's reasoning by, in any way, signalling that enough has been said, even though the examinee might seek such signals.

II.4.3 Do not endorse or criticize the examinee's fact-finding or reason-giving.

Stage Three

III.1 If "neither" is the answer chosen, then say: "So you believe neither is more believable?" Wait for examinee to respond and *then proceed to next item.*

III.2 If the item is an inference vs. observation question, that is, a question testing Principle I in Table 1, and

III.2.1 If the examinee identifies a criterion,[a] makes a comparison on the basis of that criterion, and cites a general principle, *then proceed to the next item.*

III.2.2 If the examinee identifies a criterion, makes a comparison on the basis of that criterion, but cites no general principle, then probe: "So (state the criterion mentioned) makes the difference?" with emphasis on the statement of the criterion. Wait for the examinee to respond and *then proceed to the next item.*

III.2.3 If the examinee identifies a criterion, but makes no comparison on the basis of that criterion, then probe: "Could you tell me more about the difference (state the criterion mentioned) makes to your thinking?" Wait for the examinee to respond and *then proceed to the next item.*

III.2.4 If the examinee does not identify a criterion, then probe: "Could you explain a little more what makes you believe one report more than the other?" Wait for the examinee to respond and *then proceed to the next item.*

III.3 If the item tests any principle other than Principle I, and

III.3.1 Same as III.2.1

III.3.2 Same as III.2.2

III.3.3 Same as III.2.3

III.3.4 If the examinee does not identify a criterion, then probe: "Did (state the criterion assumed when designing the item) play any part in your thinking?" If the response is affirmative, then probe: "Could you explain the part it played?" Wait for the examinee to respond and *then proceed to the next item.* If the response is negative, *then proceed to the next item.*

III.3.5 If the examinee explicitly rejects as a criterion the one assumed when designing the item, then probe: "Could you tell me some more about why (state the criterion assumed when designing the item) does not make a difference to your thinking?" Wait for the examinee to respond and *then proceed to the next item.*

III.4 Interviewing strategies for Stage Three:

III.4.1 If the examinee identifies more than one criterion including the one assumed when designing the test, then probe about the assumed criterion.

III.4.2 If the examinee identifies more than one criterion not including the one assumed when designing the test, then probe for the first one identified.

III.4.3 If there is doubt about categorizing a response, then choose the less leading probe, that is, the one that comes first on the list.

[a]A criterion is a fact or combination of facts cited in the text that can be used to make a comparison between the observers, observation conditions, or observation reports. In Item 3, the criterion is the combination of facts that Wang was involved in the accident and Vernon was not involved.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 3, September 1992, 177-188*

Judith K. Bernhard
Ryerson Polytechnical Institute

Gender-Related Attitudes and the Development of Computer Skills: A Preschool Intervention

Computer instructional programs designed to achieve equitable educational outcomes for girls have yielded mixed results. The present intervention involved 59 preschool boys and girls. Within three classrooms, separate girls' and boys' microcomputer groups were formed and matched respectively with girls' and boys' control groups (nature groups). The microcomputer groups received training in open-ended software and were taught basic skills in the LOGO language, over a six-week period. The control groups participated for the same amount of time in a variety of moderately novel, nature-related activities that were designed as gender-neutral. Both groups had female instructors and same-sex pairing. There were two major findings. The first result was that the girls' microcomputer group did not show more positive gender-role attitudes toward computers on the posttest when compared with the girls' control group. Both groups of girls viewed computers as more appropriate for their own gender than on the pretest. The second result was that in spite of equivalent instruction designed to hold the interest of girls, boys completed a significantly greater number of LOGO tasks than girls after the intervention. The question of possible initial advantages of the boys is discussed. Educational interventions with equitable goals face great difficulties in achieving and demonstrating effectiveness. This study suggests that measures of performance should be included in such demonstrations.

Les programmes d'enseignement par ordinateur conçus pour rendre équitables les résultats scolaires des filles ont connu des succès mitigés. La présente intervention a porté sur 59 garçons et filles d'âge préscolaire. Au sein de trois classes, des groupes-microprocesseurs de filles et de garçons étaient formés et appareillés, respectivement, avec des groupes de contrôle de filles et de garçons (groupes-nature). Les groupes-microprocesseurs ont reçu une formation en logiciel "ouvert" (open-ended) et ont appris les rudiments du langage LOGO, pendant six semaines. Les groupes de contrôle ont participé pendant la même durée à une variété d'activités un peu nouvelles, reliées à la nature, et conçues comme sexuellement neutres. Tous les deux groupes avaient des femmes comme instucteurs, et étaient appareillés par sexe. Il y avait deux résultats majeurs. Le premier résultat était que le groupe-microprocesseur de filles n'a pas montré des attitudes plus positives concernant le rôle sexuel à l'égard des ordinateurs au post-test, comparé au groupe de contrôle de filles. Tous les deux groupes de filles ont trouvé l'ordinateur plus approprié pour leur propre sexe qu'au pré-test. Le deuxième résultat était que, en dépit d'une instruction équivalente, faite pour maintenir l'attention des filles, les garçons ont terminé un nombre plus grand, d'une manière significative, de travaux en LOGO que les filles, après l'intervention. Les interventions éducatives avec des buts équitables connaissent de grandes difficultés dans l'atteinte et la démonstration de leur efficacité. Cette étude suggère que des mesures de performance doivent être incluses dans ces démonstrations.

Judith Bernhard is a faculty member at the School of Early Childhood Education. Her research interests include interventions to improve the learning processes and outcomes of girls and ethnoculturally diverse groups.

In fields such as science, mathematics, and computing there are strong patterns of occupational segregation whose effects include lower income, job prestige, and social status for women. The underrepresentation of women in high-level science and computing careers has been the focus of various government efforts to study the inequity and address its causes (e.g., Science Council of Canada, 1982; Status of Women Canada, 1989). Women are inadequately prepared and served by the educational system; they are less likely than men to graduate from university with computer science as a major. In 1988, 21.5% of Canadian bachelor's degrees in computer science were earned by women, as opposed to 78.5% by men (Statistics Canada, 1989). Similarly, women are underrepresented in high school mathematics and computer courses (Ontario Ministry of Education, 1988). These women are said to have "freedom of choice" and "equal opportunity," and clearly they make choices. Yet there is a pattern of avoidance determined in turn by lack of preparation at even earlier levels. Already in elementary school, gender differences in computer attitudes and performance have been consistently noted (Carmichael, Burnett, & Higginson, 1985; Hawkins, 1985). There are, in fact, gender differences in computer use and gender-related attitudes toward computers at the preschool level (Collis & Ollila, 1985; Elliott, 1989; Klinzing, 1985). In particular, girls may view use of the computer as inappropriate to their gender. This study addresses the question: What are effective instructional means of changing these attitudes and improving girls' computer skills?

A number of social factors have been cited as contributing to girls' poor performance in the computer domain. Some of these are beyond the scope of this paper: lack of confidence in computer abilities (Ogletree & Williams, 1990; Wood, Collis, Kass, Kieren, & Therrien, 1988), and girls' beliefs that technological careers are not in their future (Eccles, 1987). Other factors include teacher attitudes (Woodill, Anderson, & Bernhard, 1985), advice of school counselors (Stewart, Hutchinson, Hemingway, & Bessai, 1989), parental encouragement (Carmichael et al., 1985), the nature of the software (Malone, 1981; Turkle, 1984), group dynamics in mixed-sex groups (Guntermann & Tovar, 1987; Webb, 1984), and the ratio of children to computers (Lipinski, Nida, Shade, & Watson, 1986).

Girls' attitudes toward computers have been the focus of several interventions. Two approaches have emerged: those based on providing equal opportunities and those based on special programs for girls to ensure equitable treatment and promote equal outcomes. Collis and Ollila (1990) used an equal opportunity approach in which first graders were exposed to language arts programs. In an earlier study, Collis (1985) had found that by junior secondary school school computer use is seen as a masculine activity (see also Lockheed, 1985; Sanders, 1984). Collis and Ollila (1990) believed that the implementation of a writing program with computers would lead the girls to see that computer use was appropriate to their gender. Despite the intervention, the girls, like the boys, still saw the computer as masculine. The researchers concluded that "gender-related differences in computer use are likely to continue for students as they move through school" (p. 19). Although resistance to change is a significant phenomenon, this conclusion seems premature and unduly pessimistic; it is important to consider possible shortcomings of this unsuccessful

intervention. It is known, for instance, that boys often tend to monopolize computer terminals (Carmichael et al., 1985; Elliott, 1989; Klinzing, 1985). To achieve effectiveness, it seems important to ensure that both sexes have equal exposure time to the computer and to computer instruction.

Recently, an increasing number of interventions based on the equity approach have been published; those cited below used the LOGO computer language. Siann, MacLeod, Glissov, and Durndell (1990) developed a 12-week LOGO program for nine-year-olds. The children worked in pairs with female instructors; half of the pairs were same-sex. By assigned group, children used LOGO for either list processing or turtle-graphics activities. There was no control group with noncomputer activities. Data indicated that gender role differences in attitudes toward computers were significantly reduced regardless of pair composition or type of LOGO used. The reduction of differences was mainly due to the boys becoming much less confident in endorsing such statements as "boys are better at computers." It was noted, however, that the boys tended to "physically crowd girls out" (p. 189). Following the interventions, girls showed only a very small positive increase in response to the statement "I would be good at using computers." Further, the girls generally reported more anxiety after the computer experience. Neither boys nor girls were assessed in computer skills. Based on the disappointing results, the authors recommended that "very careful steps" (p. 190) be taken so as to avoid exacerbating gender differences in computer use. They recommended interventions which "proceed on many fronts" (p. 190); for instance, school-wide programs promoting equity (see Sanders & Stone 1986), formation of girls' computer committees, open-house programs stressing equity, and saturation of an institution with computers so as to eliminate the struggle for access.

Hawkins (1985) and her colleagues conducted a two-year intervention with 8- to 12-year olds. There were no noncomputer control groups. Having noted gender differences in the first year of their study, they attempted to make conditions more favorable to girls. The teachers adopted several new approaches and activities. Projects were designed that were thought to be of special interest to the girls; for example, programming word games. Generally, the teachers tried to spend more time with the girls. Nevertheless, by the end of the second year, gender differences (in favor of the boys) in interest and usage continued to be evident, particularly among the 11- and 12-year-olds. At all levels boys outperformed girls with respect to knowledge of programming.

There have been other LOGO interventions with preschoolers, not specifically designed to be girl-friendly, but which did address gender differences in skill acquisition. Hughes, Macleod, and Potts (1985) intervened with 17 six-year-olds over a five-month period. Children were instructed in mixed pairs. Postintervention assessment of LOGO competence indicated that there was no significant difference between the sexes. Sprigle and Schaefer's (1984) results for LOGO instruction for three- and four-year-olds were similar.

There are strong indications that gender stereotyping for computers peaks in grade 1 (Wilder, Mackie, & Cooper, 1985) and that once these attitudes and patterns are established, they become extremely resistant to change (Collis & Ollila, 1990). Hence it is clearly desirable to carry out interventions in the earliest grades before stereotypic behavior patterns are firmly set. At the time

of this writing, however, no published equity-based intervention studies with preschool children have been found. Although the above studies showed limited results of ensured equal use, it was decided that further investigation of the equal use approach was warranted. It was further decided for purposes of rigor, to use control groups engaged in noncomputer activities. The microcomputer intervention of the present study was designed especially to target preschool girls and had various special features: (a) female role models, (b) same-sex pairs, (c) software judged to be gender-neutral, and (d) ensured, equal use time. The focus of the microcomputer intervention was a controlled assessment of children's learning of the LOGO computer language.

Because there was some evidence that attitudes can be changed, it was predicted that the treatment group's gender-role stereotypic attitudes would decrease significantly in relation to those of the control group. Because the evidence has been mixed and inconclusive as to performance, the following null hypothesis was proposed: boys and girls will perform equally on LOGO skills at the posttest.

Method

Subjects

Two laboratory preschools and a metropolitan daycare center were involved. The subjects were 26 girls and 33 boys (total n=59) between the ages of 3.0 and 6.25 years. Although all subjects had some computer exposure because of the presence of a microcomputer in each classroom, they had no previous experience with the software used in the study. Due to the verbal nature of the measures, the children's receptive vocabulary level was assessed. In order to test whether the groups differed in vocabulary level, the Peabody Picture Vocabulary Test-Revised (PPVT-R, Dunn & Dunn, 1981) was administered. The boys obtained a mean score of 112.5 and the girls a mean score of 111.7. A 2x2 ANOVA with Sex (2) and Group (2) as independent variables and children's PPVT-R scores as the dependent variable revealed no significant differences between the sexes or between the treatment groups. Further, chi-square analysis of a parent questionnaire (Center for Young Children, 1983) revealed no significant differences between the sexes or the groups on computer experience outside the school, or the number of families that owned microcomputers.

Design

The study was a between-within design with three, between-subject factors (Age, Sex and Group) and one, within-subject factor (Occasion; pre- and posttest). To measure the effect of age as an independent variable, a median split by age was used to form two age groups, the first ranging from 36 months to 50 months (3.0-4.12 years; n=29), and the second ranging from 51 months to 75 months (4.25-6.25 years; n=30). Two experimental groups, a boys-only (n=17) and a girls-only (n=12) treatment group ("microcomputer groups") were formed and received instruction in microcomputers. It was not possible for any students to practice their LOGO skills in the classroom, but only during the instructional session. The two groups were matched respectively with a boys-only (n=16) and a girls-only (n=14) control group ("nature

groups") who were instructed in nature-related activities. In all cases, children were taught in same-sex pairs.

Assistants

All measures were administered by two female research assistants unaware of the experimental hypotheses and group assignment. The instructional programs were carried out by two other female research assistants (the "instructors"), also unaware of the experimental hypotheses. All research assistants had completed an undergraduate course relating to computers and young children.

Instructional Material

Microcomputer. The six-hour microcomputer package was based on EZ-LOGO and included the following open-ended software: The Friendly Computer (MECC), Sesame Street Mix and Match (Sesame Street), Facemaker (Spinnaker), Drawing: Early Games for Young Children (Springboard). The package contained a graded series of LOGO tasks beginning with the commands that are used to move the turtle forward or to its right according to its current location. The next set of activities involved exploring the length and width of the screen, directing the turtle over an acetate maze, and "teaching" the turtle to land on a sticker. The last series of tasks included various games:

1. getting the turtle to its box;
2. "teaching" the turtle to travel to each of six stickers pasted on the computer screen;
3. drawing mazes on an acetate transparency, and then "teaching" the turtle to travel through the path. The package was designed so as to permit individual instruction according to each child's progress.

Nature. These activities were designed to be either neutral or in accord with traditional classroom gender roles. Activities involved active exploration, manipulation and discovery by the child, for example, painting with ice cubes, making seed collages, and learning parts of a plant. Later activities included sorting living and nonliving things, dissolving substances, and learning about pitch by blowing on straws of different lengths.

Measures

Activity Object Instrument (AOI). Gender role stereotyping toward computer-related objects and activities was measured with the Activity Object Instrument (Collis & Ollila, 1985), an adaptation of Downing and Thompson's (1977) Activity-Object Opinion Survey, which has been validated with an international population (Downing et al., 1979). The AOI is a booklet of 16 sketches, 8 of objects without a stick figure present, and 8 of stick-figures participating in an activity. Four domains are represented in the 16 items included in the AOI: Four of the sketches are reading-related, four writing-related, four microcomputer-related, and a fourth category includes miscellaneous items. Subjects are presented with a sketch of an object or activity and required to decide if a boy or a girl would be more suitable in the picture. Each domain is scored as follows: A score of 2.00 means that "boy" and "girl" were chosen equally. A score greater than 2.00 indicates that "boy" was selected more often than "girl," and a score below 2.00 that "girl" was selected more often than "boy." The measure for each domain has a range of 0-4. Validity

and reliability of the AOI are reported to be adequate (Collis & Ollila, 1990). In the present study, only scores for the microcomputer domain were used.

LOGO skills tasks. The children's computer skills were measured with the LOGO Skills Tasks (LST), an informal, individually administered series of three tasks developed for the present study. Tasks were to be completed by each child without time limits. Each of the three tasks began with the cursor in the middle of the screen. An initial practice task to ensure comprehension involved asking the child to take the turtle to several places on the screen. The first task consisted of making the turtle go from the center of the screen to the upper, right-hand corner of the screen ("Turtle to sticker"). The second comprised joining the stars at the four corners of the screen ("Joining four stickers"). The third and final task involved leading the turtle through a maze ("Maze"). Each child was assigned a score of 1 or 0 on each task according to whether they were successful or not; the efficiency of their path was not considered. For an overall measure of computer skill, the three task scores were summed. Therefore, this measure had a range of 0 to 3.

Peabody Picture Vocabulary Test—Revised. In order to control for the effects of receptive vocabulary, the Peabody Picture Vocabulary Test—Revised (PPVT-R; Dunn & Dunn, 1981) was administered.

Parent Questionnaire (PQ). The parent questionnaire developed by the Center for Young Children (1983) was designed to elicit information regarding children's exposure to, and use of, electronic media outside the school setting. There were also questions regarding parental views and expectations of children's use of computer technologies. For example, one set of questions asked, "Outside the school does your child have experience with microcomputers, video games, and/or electronic aids?" Another set of questions asked whether there was a computer in the home, who used it, and how often. A third set of questions enquired into parental views of advantages and disadvantages of young children using computers at school. The data from the parent questionnaire were tabulated and analyzed by performing a chi-square analysis in order to determine significant differences between groups or genders.

Procedure

The pretests (AOI, PPVT-R, PQ) were administered three weeks before treatment, and the posttests (AOI, LST) were administered within the week following the termination of the intervention. All test measures were administered by two female research assistants. Tests were individually administered.

Treatment. The microcomputer and nature instructional programs were both carried out over a six-week period; both programs comprised 18 sessions, each of which lasted approximately 15-20 minutes. A female research assistant took the children to a quiet room in the school and worked with them in same-sex pairs.

Microcomputer. The open-ended software was presented during the first two weeks in order to familiarize the children with the names of parts of a computer and the proper handling and loading of disks. After being introduced to basic LOGO commands, children were taught the strategy of pretending to be a turtle as a means of clarifying direction commands. In later sessions involving turtle activities, the children were encouraged to explore

and were verbally reinforced for their efforts. Every effort was made to ensure the child's interest in and understanding of the tasks.

Nature. The instructor presented the materials and suggested the activities that could be carried out with the paint, ice cubes and paper. As the children proceeded, the experimenter made suggestions or comments designed to elicit a variety of exploratory activities. Every attempt was made to get the children interested and engaged in the activities.

Results

Gender Role Attitudes Toward Computers

Children's pretest and posttest scores on the computer domain of the AOI are shown in Table 1. The score indicates the number of times (out of four) that "boy" was selected as more suitable for a presented sketch. A score of 2.00 represents a neutral response, whereas a score less than 2.00 indicates that "girl" was chosen more times than was "boy." Neither the boys' nor the girls' microcomputer group compared with their respective control changed in their gender role attitudes; that is, their view of whether the computer was appropriate to their gender. A significant main effect for Sex was found [$F(1,51)=106.6$, $p<.001$]. Because the boys' responses (mean=3.31) were further from the neutral of 2.00 than the girls' responses (mean=1.33), the boys showed more gender stereotypic attitudes toward computers than did the girls. There was also a significant Sex-by-Occasion interaction [$F(1,51)=10.24$, $p<.002$] indicating that both boys' and girls' tendency to choose own-sex responses increased across the two occasions. No age differences were found.

Computer (LOGO) Skills

Table 2 shows mean scores for the LOGO Skills Tasks (LST). The boys completed a mean of 2.47 tasks (SD=.62) and the girls completed a mean of 1.58 tasks (SD=1.0). In order to determine if this difference was significant, a Mann-Whitney Rank Sum was computed and found to be significant [$U=49.5$, $p<.02$]. Thus, in spite of equivalent instruction, the boys completed a significantly greater number of tasks than did girls.

In order to obtain more detailed information on the pattern of results across the different computer tasks, a chi-square analysis test (with Yates correction) was performed (see Table 2). The girls consistently completed fewer tasks than

Table 1
Means and Standard Deviations for Gender-Typing of Computer Use

Subjects	*Pretest*	*Posttest*
Boys	3.06 (0.99)	3.55 (0.83)
Girls	1.62 (0.85)	1.04 (1.08)
Microcomputer	2.40 (1.16)	2.50 (1.59)
Nature	2.45 (1.21)	2.38 (1.57)
Young	2.44 (0.99)	2.53 (1.58)
Old	2.39 (0.99)	2.30 (1.58)

Table 2
Chi Square Analysis on LOGO Skills Tasks (LST) by Gender

LOGO Task	*Males*	*Females*	*Chi-Square*
1. Turtle to sticker	17/17	9/12	2.43
2. Joining four stickers	15/17	5/12	5.12*
3. Maze	10/17	5/12	0.28

*significant at $p<.05$

the boys, but only one of the three tasks considered alone showed a significant difference, the "Joining four stickers" task).

In order to determine if there were significant age differences in ability to complete the LST, a Mann-Whitney Rank Sum was computed, and it was found that the older children completed a significantly greater number of tasks than the younger children [U=35.5, $p<.002$]. A follow-up chi-square analysis (with Yates correction) was applied in order to consider the tasks separately. Table 3 shows the pattern of age-related task performance: the older children completed significantly more tasks than the younger children.

Qualitative Observations

The superior performance of the boys was paralleled by their greater eagerness to learn LOGO. There were several obvious differences between the behavior of boys and girls during the intervention. Girls generally showed a desire to finish the task satisfactorily and sought immediately to return to the classroom, although they had apparently liked the tasks. During teaching they awaited instructions and asked permission to do things. Boys, on the other hand, seemed to find the tasks more interesting and explored eagerly. They raised questions about procedural matters. They shared among themselves to a greater degree than did the girls. Boys often tended to remain at the computer after their tasks were completed. In sum, the boys acted more avidly and inquisitively than did the girls in this computer domain. This difference may have existed in noncomputer domains as well, but no data were gathered that could bear on this issue.

Table 3
Chi Square Analysis on LOGO Skills Tasks (LST) by Age

LOGO Task	*Old*	*Young*	*Chi-Square*
1. Turtle to sticker	14/15	12/14	0.004
2. Joining four stickers	13/15	7/14	2.99
3. Maze	12/15	3/14	7.74*

*significant at $p<.05$

Discussion

Issues Related to Attitudes

Girls. At the posttest, the girls' microcomputer group did not differ from their control regarding perception of the appropriateness of computers for girls. Both groups of girls, regardless of treatment, began by viewing computers as gender appropriate, and both moved in the desired direction of perceiving computers as even more gender appropriate (i.e., "for girls"). The content, that is, subject matter, of the two programs in the present study did not appear to be a major factor in producing changes in the girls' attitudes. It may therefore be argued that, in the absence of maturational changes, the factors common to the two programs may have influenced the girls' attitudes in both groups. (It would seem unlikely that significant maturational changes occurred in the six-week period.) Before discussing what the common elements might be, the present findings are related to those already reviewed. Siann et al. (1990), with female instructors and other girl-friendly measures, reported a reduction of gender differences in attitudes. In view of the present findings, it is entirely possible that the causes of the shift were not connected with the computer-related aspects of the intervention. In attempting to ascertain the effective noncomputer elements of Siann et al., we return to the issue of common elements of the two treatments in the present study: there was an "investigative curriculum"; it was presented in conditions favorable to girls (same-sex pairs, female role models). These are the most obvious elements that may have helped girls in both groups feel more positive about their role in computing. It is also possible that the girls in the microcomputer group talked to the other girls and influenced them with respect to computer-related attitudes. This possibility is somewhat diminished because the microcomputer girls could not practice newly learned skills at the regular classroom computers used by other girls.

In view of the lack of specific results, the direction of Siann et al.'s (1990) comments seems particularly appropriate. They did not suggest abandoning programs, but rather instituting more comprehensive approaches. Approaches with many fronts appear to be necessary to effect changes in educational outcome. Such approaches might include such measures as unequal amounts of time in the girls' favor, special compensatory instruction, social use of computers, parent workshops, and classroom visits by women in traditionally masculine occupations.

Boys. Siann et al. (1990) found that the boys' group after intervention were less likely to agree with statements about boys' superior computer abilities. Hence the question arose for the present intervention, Would the boys' microcomputer group become less stereotyped? Would this group as compared with its control come to believe that the computer was more appropriate to girls (i.e., less exclusively "for boys")? In fact, the boys' microcomputer group showed no such difference compared with their control group. At the posttest, the boys of both groups had undergone rapid changes and believed *more strongly* than at the pretest that computers were for their own gender. Thus, the boys' attitudes were quite sensitive to the experimental manipulation. Of course, all the boys in the present study were in same-sex pairs. Nevertheless, apart from the factor of pairing, the general question remains

unanswered: Are boys, when confronted with an instructional program involving females' increased presence, direction, and modeling likely to become less or more gender stereotyped in their attitudes toward computers?

Issues Related to Skill Acquisition

According to hypothesis B, it was predicted that the boys and girls would perform equally on LOGO-related tasks after the microcomputer instruction. The hypothesis was not accepted. On the LOGO Skills Tasks (LST) administered to the microcomputer group it was found that there were sex differences: boys completed significantly more tasks than girls.

With the benefit of hindsight, one can infer that the features in the computer instructional program intended to benefit the girls were probably insufficient to overcome possible initial advantages. The nature and source of boys' advantages in computer programs as well as math and science programs is extensively documented (see e.g., Harding, 1983, Sanders & Stone, 1986). Advantages may come from (a) initial skills, (b) stronger interest and motivation, (c) gender differences in play styles, and (d) gender differences in sharing and collaborative work.

Boys may have had superior initial skills because of the situation of the computer in their families. Although the data showed no significant gender differences regarding the presence of microcomputers in the home, this did not rule out all gender differences in the amount and manner of usage of microcomputers in the home. Elliott (1989), for instance, found that although 20% of all children had a computer in the home, 70% of boys had used it, compared with only 30% of girls. Furthermore, the boys in Elliott's study had much more knowledge of computer functioning than girls and were more familiar with the terminology relating to computers.

Regardless of the specific source of boys' advantages in the present study, their superior performance is consistent with the findings of Hawkins (1985). The present qualitative observations generally support one of Hawkins' suggestions as to the cause of girls' lower achievement. She labeled this factor *interest*, and it is clearly the same factor that has been here described as eagerness, enthusiasm, and avidity.

Successful performance is a powerful motivator for the development of positive attitudes and expectations. Where performance is lower, an individual may gradually lose motivation, especially where lower performance is consistent with gender-role stereotypes. It seems clear then, that interventions should not merely be aimed at changing gender-role attitudes but also at improving performance. The effect of the present intervention on skills but not attitudes in the microcomputer groups suggests the importance of performance measures in future, equity-based interventions.

Implications for Education

The postintervention gender differences in performance found at this early age indicate that similar conditions of instruction, even girl-friendly approaches, may not be sufficient to change girls' stereotyped perceptions and to impart skills comparable to those of boys. Equitable programs (i.e., those sufficient to achieve equality of outcomes) must specifically address the disadvantages of girls that already exist at the preschool level and do not simply

appear at puberty. These disadvantages need to be dealt with long before the beginning of high school. More comprehensive equity programs are needed to bring girls into the fields of mathematics, science, and computing so that they consider these domains as entirely appropriate to their gender and vital to their careers.

Acknowledgments

This research could not have been carried out without the cooperation of the children and teachers of the Early Learning, Hydrokids, and George Brown centers. Preparation of this article was supported in part by a grant from the Ryerson Polytechnical Institute Faculty of Community Services Research and Scholarship Committee and a doctoral fellowship from the Social Sciences and Humanities Research Council of Canada. Apple Canada graciously supplied the microcomputers. Linda Siegel, Peter Lindsay, Paula Caplan, Sherry Turkle, and Betty Collis provided valuable suggestions. The editorial assistance of Hal White is gratefully acknowledged. Special thanks to Nancy Hutchinson and Robert Short for their support in revision for publication.

References

Carmichael, H., Burnett, D., & Higginson , W. (1985). *Computers children and classrooms: A multisite evaluation of the creative uses of microcomputers by elementary school children.* Kingston, ON: Ontario Ministry of Education.

Center for Young Children. (1983). *Parent computer questionnaire.* Unpublished questionnaire, College Park, MD: University of Maryland.

Collis, B. (1985). Sex-related differences in attitudes toward computers: Implications for counselors. *The School Counselor, 33*(2), 120-130.

Collis, B., & Ollila, L. (1985). *Gender stereotypes about reading, writing, and computer use in preschool, kindergarten, and grade 1 boys and girls.* Unpublished manuscript, University of Victoria.

Collis, B., & Ollila, L. (1990). The effect of computer use on grade 1 children's gender-stereotypes about reading, writing, and computer use. *Journal of Research and Development in Education, 24*(1), 14-20.

Downing, J., Dwyer, C.A., Feitelson, D., Jansen, M., Demppainen, R., Matihaldi, H., Reggi, D.R., Sakamoto, T., Taylor, H., Thackray, D.V., & Thomson, D. (1979). A cross-national survey of cultural expectations and sex-role standards in reading. *Journal of Research in Reading, 2*(1), 8-23.

Downing, J., & Thomson, D. (1977). Sex-role stereotypes in learning to read. *Research in the Teaching of English, 11*, 149-155.

Dunn, I.M., & Dunn, L.M. (1981). *Peabody picture vocabulary test—Revised.* American Guidance Survey.

Eccles, J.S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly, 11*, 135-172.

Elliott, A. (1989). Girls, computers, and early childhood education. In N.J. Kyle (Ed.), *Women as educators in 19th and 20th century Australia.* Wollongong: University of Wollongong.

Guntermann, E., & Tovar, M. (1987). Collaborative problem-solving with LOGO: Effects of group size and group composition. *Journal of Educational Computing Research, 3*(3), 313-334.

Harding, J. (1983). *Switched off: The science education of girls.* York, UK: Longman for Schools Council.

Hawkins, J. (1985). Computers and girls: Rethinking the issues. *Sex Roles, 13*(3/4), 165-180.

Hughes, M., Macleod, H., & Potts, C. (1985). Using LOGO with infant school children. *Educational Psychology, 5*(3-4), 287-301.

Klinzing, D.G. (1985). *A study of the behavior of children in a classroom equipped with computers.* (ERIC Document Reproduction Service No. ED 255 320)

Lipinski, J.M., Nida, R.E., Shade, D.D., & Watson, J.A. (1986). The effects of microcomputers on young children: An examination of free-play choices, sex differences, and social interactions. *Journal of Educational Computing Research, 2*(2), 147-166.

Lockheed, M.E. (1985). Women, girls, and computers: A first look at the evidence. *Sex Roles, 13*(3/4), 115-122.

Malone, T. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science, 4*, 333-369.

Ogletree, S.M., & Williams, S.W. (1990). Sex and sex typing effects on computer attitudes and aptitude. *Sex Roles, 23*(11/12), 703-712.

Ontario Ministry of Education (1988). *September school report forms,* Curriculum System, Table 11.47P, unpublished data.

Sanders, J.S. (1984). The computer: Male, female or androgynous? *The Computing Teacher, 1*(18), 31-34.

Sanders, J.S., & Stone, A. (1986). *The neuter computer: Computers for boys and girls.* New York: Neil-Schuman.

Science Council of Canada (1982). Who turns the wheel? *Proceedings of a workshop on the science education of women in Canada.* Ottawa: Author.

Siann, G., Macleod, H., Glissov, P., & Durndell, A. (1990). The effect of computer use on gender differences in attitudes to computers. *Computers in Education, 14*(2), 183-191.

Sprigle, J.E., & Schaefer, L. (1984). Age, gender, and spatial knowledge influences on preschoolers' computer programming ability. *Early Child Development and Care, 14,* 243-250.

Statistics Canada. (1989). *Science and technology economic analysis.* Policy sector, industry, science, and technology, unpublished data.

Status of Women Canada. (1989). *Participation of girls and women in math, science and technology.* Paper submitted to the joint meeting of federal/provincial/territorial ministers responsible for the status of women and provincial/territorial ministers responsible for education.

Stewart, H.M., Hutchinson, N.L., Hemingway, P., & Bessai, F. (1989). Program advice given to Canadian adolescents: Effects of student gender, ethnicity, and achievement. *Alberta Journal of Educational Research, 35*(3), 255-266.

Turkle, S. (1984). Women and computer programming. *Technology Review,* November/December, 48-50.

Webb, N. (1984). Sex differences in interaction and achievement in cooperative small groups. *Journal of Educational Psychology, 76*(1), 33-44.

Wilder, G., Mackie, D., & Cooper, J. (1985). Gender and computers: Two surveys of computer-related attitudes. *Sex Roles, 13*(3/4), 215-228.

Wood, P.H., Collis, B., Kass, H., Kieren, T., & Therrien, D. (1988). *Gender-related access patterns and types of computer usage among Canadian adolescents.* Paper presented at the Canadian Society for Studies in Education, Windsor, Canada.

Woodill, G., Anderson, K, & Bernhard, J. (1985). *Sexism, computers, and young children.* Paper presented at the joint meeting of the Association of Educational Data Systems and the Educational Computing Organization of Ontario, Toronto.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 3, September 1992, 189-204*

Ellen M. Regan
OISE Department of Applied Psychology

and

John A. Ross
OISE Trent Valley Centre

Professional Development for School Consultants: An Evaluation[1]

Most empirical studies reporting effects of consultant training reveal such flaws as inadequately described treatments, use of vague self-report measures, and no comparison/control group. This article describes the effects of a consultant inservice that was focused on specific consulting strategies derived from research on consultant effectiveness and delivered through a program designed according to a "reflective" consultant role orientation. The inservice was assessed through pre- and posttreatment interviews that revealed strategies used by consultants when working with individuals and groups. In the first year of the study, the performance of 18 treatment consultants was compared with 18 matched controls. In the second year the program was revised and delivered to 35 consultants. In the first year the treatment group modestly outperformed the controls; the second year treatment consultants learned substantially more than both groups. The impact of the program was attributed to the use of professional case studies as stimulus for reflection. Consultants are professionals who help other professionals achieve job related objectives. Educational consultants help teachers and other staff implement school change and improve practice. For example, in one Canadian province the function of consultants is to "assist teachers in maintaining proper standards and improving instruction" (Ontario, 1984). Recent events such as reductions in the length of consultants' tenure and in the number of funded positions suggest substantial dissatisfaction with performance. Attempts to improve the quality of educational consultation have tended to focus on selection and supervision strategies, with disappointing results (e.g., Louis & Kell, 1981). The purpose of the research reported here was to develop, implement, and evaluate an inservice approach to enhancing consultant effectiveness.

La plupart des études empiriques indiquant des effets de la formation à la consultation révèlent des défauts tels que des traitements insuffisamment décrits, l'emploi de rapports vagues, et l'absence d'un groupe de comparaison/groupe de contrôle. Cet article décrit les effets d'une formation continue à la consultation axée sur des stratégies spécifiques de consultation s'inspirant de la recherche sur l'efficacité à la consultation, et délivrée à travers un programme fait à partir d'une orientation "réflexive" du rôle du consultant. La formation était évaluée à partir d'entrevues avant et après le traitement dans lesquelles étaient mentionnées les stratégies utilisées par les consultants travaillant avec des individus et des groupes. Pendant la première année de l'étude, la performance de 18 consultants dans un groupe de traitement était comparée à celle de 18 consultants dans un groupe de contrôle. Pendant la deuxième année, le programme était révisé et délivré à 35 consultants. Durant la

Ellen Regan is a professor in the Department of Applied Psychology. Her research interests include practitioner beliefs related to professional roles and the professional development of teachers and consultants.
John Ross is a professor in the Department of Curriculum and Head of the Trent Valley Centre. His research interests include instructional improvement and professional development.

première année, la performance du groupe de traitement était légèrement supérieure à celle du groupe de contrôle; pendant la deuxième année, la performance des consultants du groupe de traitement a dépassé de loin la performance des deux groupes. L'effet du programme était attribué à l'emploi d'études de cas professionnelles comme objet de réflexion.

Few studies of consultant inservice (or preservice) have been reported in the literature. Most articles merely provide a rationale for offering consultant training (e.g., Padfield, 1981) or prescribe its characteristics without assessing effects (e.g., Froehle, 1978). The few empirical studies that have been reported are flawed. For example, Broskowski, Khajavi and Mehryar (1973) claimed positive effects for a 12-week training program but the nature of the treatment was only briefly sketched; the effects were assessed by vague self-report measures and there was no comparison group. Goodwin and Coates (1973) provided lectures on theory with simulations and guided practice in lab and school contexts. They found large changes in consultants' behavior, which were associated with effects on teachers and students, but again there was no control group.

Inservice Approach

Inservice Goals

In designing the inservice, primary attention was given to identifying the competencies to be acquired. Although most studies of consultation have reported positive findings (see, for example, the review of 29 school consultation studies in Medway, 1979), only a few investigators have attempted to identify specific consultant behaviors that have an impact. Most inquiries examined the effects on teachers of variations in the *quantity* of consultation. For example, two studies (Loucks, 1982 and Louis, 1981) found that schools where support was provided by two groups of consultants (internal and external) showed greater evidence of change than schools where only one type was involved. Several other studies (Curtis & Watson, 1980; Miles, Fullan, & Taylor, 1978; Runckell & Bell, 1976; Tyler & Fine, 1974) found that the duration of consultant-teacher contact had a positive impact on teacher growth. Other investigators made finer distinctions in terms of the time variable. Loadman and Mahan (1972) found that effective consultants spent more time on inservice and modeling activities, especially demonstration teaching. Louis and Sieber (1979) found that *when* the consultant time was delivered (during implementation rather than up front) mattered more than the total amount of time given.

Few attempts have been made to identify differences in the effects of strategies used by consultants. Firestone and Corbett (1981) were not able to identify any consultant activities that consistently produced change; similar findings were reported by Louis and Kell (1981). Miles, Saxl, and Lieberman (1988) attempted to distinguish between the strategies used by external consultants who were rated as outstanding by project managers and those who were rated average: the differences were small and inconsistent.

In contrast, Ross and Regan (1990) found substantial differences in the strategies used by an exemplary group of experienced consultants compared with those used by inexperienced consultants. They conducted a series of semistructured interviews with 12 curriculum consultants (half inexperienced,

half experienced) from two large school boards, six senior administrators from the same boards, and six additional consultants from a third board with significantly different characteristics. Categories of analysis evolved from the data and were organized in a four-phase model of consultation: initiation, planning, delivery, and follow up. Although similarities in beliefs about the consulting process were found, consultant groups differed strikingly in their methods for helping teachers. These differences were interpreted in the expert-novice framework that has been used for other educational roles (e.g., Bereiter & Scardamalia, 1986; Berliner, 1986; Leithwood & Stager, 1987). The delivery and evaluation strategies used by relative experts suggested a reflective approach to consultation not demonstrated by the novices. The strategies of more experienced "reflective" consultants revealed a problem solving approach in responding to consultee requests and in consultant initiated activity. Less experienced consultants tended to accept consultee definitions of problems and to respond to such requests with pat answers or solutions.

The findings from the Ross and Regan (1990) study were used to construct a profile of consultant development containing 11 dimensions with a hierarchy of strategies for each. The four levels in each dimension represented incremental steps toward effectiveness. Ten of these dimensions and the anchoring points (upper and lower levels) on each were derived directly from the expert-novice study data. The intermediary levels were derived from the same source or were interpolated by the researchers. In all instances, the levels were validated by reviewing the school change literature.

For example, for Profile Dimension 3: *Determining the needs to be met,* the expert-novice data revealed that inexperienced consultants tended to impose their own personal agendas (Level 1) in planning activities to meet needs. One inexperienced consultant reported, for example, that his workshops replicated a provincial inservice program with specific prescribed objectives, that is, "it's quite a program ... very prescriptive ... very tight." Other inexperienced consultants described objectives for their consulting activities as deriving from personal agendas or their general knowledge of teachers in their boards. The researchers added to these data the observation that novice consultants frequently impose school system priorities without regard for the perceived needs of teachers (Level 2). In contrast, experienced consultants were concerned with responding to teacher expectations. For example, one experienced consultant referred to the importance of obtaining accurate data on teacher needs (Level 3 strategy). This consultant referred specifically to a situation in which she talked with both the teacher and principal "beforehand ... that way everybody is brought on board ... and you can help the teacher a little bit more because you have more information." Experienced consultants talked about determining needs as a process of negotiating expressed needs and priorities (Level 4). For example, one experienced consultant observed, "there's a (system) plan ... you try to get them to plug into that plan." Another experienced consultant talked about using the system plan and "data we had collected from people in their brainstorming sessions."

School change research concerned with the identification of adult learner needs provides support for the nature and sequence of all four levels on this dimension. In the literature on techniques for identifying needs of adult

learners (e.g., Coffing & Hutchison, 1974; Griffin, 1978), there is agreement that appropriate procedures consist of identifying a desired state, collecting accurate information about current status, and measuring discrepancies between the two. The three elements of this shared procedure are increasingly prominent in higher levels of the consultant profile. Imposition of the consultant's agenda (Level 1) involves none of these elements. Imposition of system priorities (Level 2) involves one element, specification of desired state. Response to expressed needs (Level 3) involves two elements—actual status and perceived discrepancies. The (Level 4) strategy, negotiation of expressed needs and system priorities, involves all three elements.

The evidence used in constructing the four strategy levels for the first 10 profile dimensions is reviewed in Regan and Ross (1991). The 11th profile dimension, *reflecting within the role*, was added later as a summative orientation to the role. It resulted from observing how consultants responded to inservice activities. This dimension distinguished between consultants who appear to be unreflective about their work and those who consciously analyze activities and outcomes on a regular basis with a view to self-improvement. The resulting growth continuum (shown in Figure 1 as the Consultant Profile) describes a gap between actual and ideal performance that was to be overcome through the inservice.

Inservice Design

The consultant inservice was founded on three principles discussed and supported in the literature as guides for helping professionals improve their practice. The principles were:

1. focus on specific strategies and skills because training that combines knowledge acquisition with heavy emphasis on skill development is potentially more effective (Goodwin & Coates, 1973; Griffin & Barnes, 1986; Miles et al., 1988);
2. provide opportunities to reflect on one's own practice (Clark, 1988; Schön, 1987);
3. provide opportunities to practice skills in framing problems and in experimenting with these frames in actual practice environments (Schön, 1983, 1987).

We saw the strategies of expert consultants as actions that could be taught to less experienced consultants by providing tasks that linked the findings from research to the consultant's own experience. The tasks that were designed collectively approximated the Gliessman, Grillo, and Archer (1989) stages of problem solving cited by Kagan (1990) in her discussion of research related to teacher self-reflection. Consultants were engaged in analyzing problems, generating problem solutions, interpreting solutions in the light of additional information, and considering the consequences of different courses of action. Kagan's review found that verbalizing one's reflections in writing or sharing with another is important to the process of self-reflection. Accordingly, the inservice tasks and materials provided for recording individual and group reflection as well as the opportunity to share reflections and problem solving strategies through group discussion.

In accordance with these principles, the inservice was designed to engage participants in specific tasks within a continuing cycle of reflection-planning-

Dimension	Level 1	Level 2	Level 3	Level 4
Before consultation begins				
1. Identifying the recipients of consultation	individual teacher	teacher in context of other teachers)	teachers and administrative team in the school	teachers and administrators in adjacent schools
2. Recruiting the consultation team	consultant works alone	consultant teams with one other agent (e.g., another consultant)	consultant teams with two other agents (e.g., consultants and expert teachers)	consultant teams with three other agents (e.g., expert teachers, school administrators, and other consultants)
Planning activities to meet needs				
3. Determining the needs to be met	imposition of consultant's personal agenda	imposition of system priorities	(a) response to expressed needs of teachers based on limited data (b) response to expressed needs of teachers identified systematic data search	negotiation of expressed needs and system priorities
4. Individualizing within the consultation	treat all teachers the same	plan extra follow-up for enthusiastic teachers	provide options for all teachers to select from	assign each teacher to most appropriate option
5. Providing practice within consultation sessions	no planned practice	teachers discuss teaching	(a) teachers experience activities as students (b) teachers simulate teaching (e.g., each other or plan units)	teachers teach children
6. Providing practice between consultation sessions	no planned practice	practice with no planned feedback	practice with feedback	provide support that enables teachers to design self-directed practice and self-appraisal
7. Using in-school expertise	no planned teacher sharing	sharing encouraged but not scheduled	planned sharing by expert teachers to spread ideas to others	planned sharing by all teachers to consolidate new ideas
Evaluation and follow-up				
8. Response to teacher feedback during delivery	ignores feedback	expands or contracts treatment of objective	shows how consultation meets teacher needs	shifts to a different system objective that better meets teacher needs
9. Identifying types of data to assess consultation	perceived satisfaction of the teacher	perceived usefulness to the teacher	objective indicators of teacher change	objective indicators of student change
10. Using evaluation data	no use	planning subsequent sessions with the same teacher	planning subsequent sessions with similar teachers	refinement of consultation strategies
11. Reflecting within the role	no reflection	reflection on incidents	reflection on specific instances within a framework	systematic experimentation to develop new strategies

References to the *teacher* in the profile should be read as the *consultee.*

Figure 1. Consultant profile.

practice-reflection. The target skills in these tasks were the specific strategies identified in the Consultant Profile. Eight case studies developed by the research team provided the anchor for the inservice. The cases were organized in four sets of two. Each set described a problem commonly experienced by consultants; one case presented the problem as experienced in a one-to-one consultation and the other presented the same problem in a group setting. Both cases in a given set addressed the same profile dimensions and reflected a Level 1 strategy for each dimension. Collectively the eight cases addressed the first 10 dimensions of the Consultant Profile. Each session was structured around Case Study Analysis, Action Planning, and, after on-site implementation of the plan, Feedback.

In Case Study Analysis there was a brief period of consultant personal reflection followed by a longer period of group sharing and analysis. Working in groups of three to four, consultants identified the strategies described in the pair of cases and their practice implications. Information on the strategies of effective consultants, based on earlier research and illustrated in the Profile, was then provided by the research team. The group leader's role was to clarify Profile and task dimensions as needed, keep the discussion on task, encourage sharing by all group members, and help the group achieve closure regarding its task. Members of the research team served as team leaders in the first year of implementation; they were replaced in the second year by consultants.

The Action Planning tasks involved analysis of a self-identified "personal case in the past" and development of a future action plan based on perceived discrepancies between current practice and the higher levels of the Consultant Profile. In addition to describing strategies to be implemented in a future consultation, consultants elaborated their plans by supplying details of how they would proceed, what they saw as alternative courses of action, and how they planned to assess the effects of their new strategies. The group leader's role in support of these activities included clarifying the task and providing feedback on the action plans, for example, encouraging precision and specificity in the plans and suggesting alternatives to be considered. Between sessions action plans were implemented by consultants in their everyday roles over a five- to seven-week period.

The Feedback tasks provided consultants with the opportunity to share and critique the implementation of their action plans. Each group was responsible for identifying the important practice implications of the strategies they used. The group leader's role in this activity was to ensure that all group members shared/critiqued their experiences and were able to identify their own growth.

Implementation and Evaluation of the Professional Development Program

The program was piloted in a truncated form in two school boards. Fourteen consultants in one board and 18 in the other analyzed one pair of case studies, reflected on their current practice, designed a plan to improve that practice, and received feedback on the implementation of their plans. Anecdotal evidence provided by participants suggested positive outcomes, but no systematic data were collected. The full program was implemented and evaluated with two different groups of consultants over a two-year period (1988-1989, 1989-1990).

The specific questions to be investigated in the research reported here were:

1. Did the inservice have a beneficial impact on the strategies used by consultants?
2. Did revisions to the inservice program in the second year increase its impact?

Study Year 1 (1988-1989)

Sample/Procedure

Eight consultants in one large board in central Ontario and 10 consultants in a medium-sized board located 125 km from the first volunteered to be the treatment sample. Consultants were given release time to attend sessions in their respective board offices. The treatment sample was matched in terms of experience and subject area responsibility with 18 volunteer consultants from adjacent boards.

In September 1988 individual interviews were conducted with subjects in the two samples. Treatment sessions (three full days and two half days) began in late October and continued until late April 1989. Post interviews for both treatment and control groups were conducted in May 1989. The Interview Guide asked subjects to describe two of their recent consultations, one in which they were working with individuals and the other with groups. Subsequent questions probed the specific strategies consultants used. The interviews were audiotaped and verbatim transcripts were made. The sample of consultant strategies described were coded as levels of the Consultant Profile. Each transcript was independently coded by two trained research assistants; disagreements between coders were resolved through discussion.

Univariate statistics for pre and post scores were generated. Because pretest interviews correlated with post scores, standardized residuals were produced by regressing post over pre scores. Differences between treatment and control groups were determined by conducting one-way analyses of variance on the residuals.[2]

Results

A substantial amount of data was missing on three Profile dimensions. "Reflection within the role" had not been included in the original version of the interview guide because our appreciation of this variable did not emerge until near the end of the inservice. Attempts to recapture performance on this dimension by examining responses to other items on the pretest administration failed. Because a pretest score could be confidently determined for only a few members of the sample, the item was dropped from the analysis. There was also a problem with "using assessment data." On the pretest several respondents in both samples misunderstood the question; the problem was resolved on the post administration by adding further probes. Finally, several consultants experienced difficulty responding to the dimension concerned with "teacher feedback during delivery of consultation." Some persisted in describing actions that would be taken at a subsequent consultation (rather than during the session in which the problem arose); others indicated they were unable to answer because they had never encountered negative feedback during delivery. These responses, almost half the total for the teacher feedback items in both administrations, were coded as missing.

Table 1 displays the means and standard deviations of each profile dimension for each group on each test occasion; the table also shows the average based on 10 dimensions. (The average was used as the summative outcome measure rather than the total because of missing data.) Table 1 indicates that the means increased slightly over the duration of the project in both samples. The analyses of variance in the residuals found that the treatment group significantly outperformed the control group overall [$F(1,34) = 4.22$, $p<.048$], although the differences on the individual dimensions reached significance only for determining needs [$F(1,34) = 5.35$, $p<.027$] and response to teacher feedback [$F(1,8) = 7.57$, $p<.025$]. The overall effect size of the program using the formula of Glass, McGraw, and Smith (1981) was a modest .26.

Discussion

Consultants enrolled in the inservice were effusive in their appraisals of the program. They perceived that it had a large effect on their practice, far more than was indicated by the measured outcome. The most likely explanation for the participants' enthusiasm is that the program filled a void: there were no competing programs that the consultants could compare it with and it provided an otherwise rare opportunity for consultants to talk to peers about their work. But why did the inservice have such a small effect on what consultants do? Several explanations came to mind.

1. The upper levels of the Profile were developed by selecting the best efforts of an outstanding group of consultants, and the definition of these levels was further enhanced by incorporating findings from research on school change. The difficulty in moving from one level of the Profile to another may have been underestimated. In several of the sessions the inservice designers attempted to address several different components of a single dimension or combined several complex strategies. Some participants may have experienced overload, particularly since the inservice did not set priorities among the dimensions and did not attempt to provide different levels of inservice corresponding to the specific needs of each of the candidates. In addition, there may have been insufficient practice to consolidate the new levels of performance into existing practice. Each dimension was addressed only once; no further opportunity for practice was given, and there was no attempt to accumulate the dimensions in successive action plans.
2. The academics who delivered the inservice may have been too active in the deliberations of the small groups. Observations by research team members during the inservice sessions found that comments/questions tended to be addressed to the research team members rather than to fellow consultants, thus influencing the nature of group interaction and the topics of discussion.
3. Responses to certain tasks and materials suggested a need for revision to highlight particular strategies, clarify task requirements, and modify procedures. For example, there was confusion about the meaning of some dimensions, specifically reacting to teacher feedback during delivery of consultation and individualization within consultation. Because attention to these consulting dimensions is so dependent on consultant on-the-spot awareness and sensitivity, an alternate means of treating these dimensions

Table 1
Means and Standard Deviations for 1989 Study Outcome Measures

Profile Dimension	Pretest						Posttest					
	Treatment			Control			Treatment			Control		
	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD	N	$\bar{X}$	SD
Identifying consultation recipients	18	1.69	0.30	18	1.58	0.19	18	1.86	0.51	18	1.61	0.21
Recruiting consultation team	17	1.79	0.69	18	1.39	0.63	17	2.24	0.73	18	1.61	0.61
Determining needs	18	2.69	0.42	18	2.61	0.45	18	3.07	0.48	18	2.65	0.56
Individualizing within consultation	18	1.72	0.96	16	1.81	1.17	18	1.94	1.16	17	1.76	1.03
Practice within sessions	16	2.56	1.00	17	2.31	0.98	17	2.60	0.78	18	2.49	0.90
Practice between sessions	13	2.04	1.01	15	1.90	0.63	16	2.13	0.63	17	2.32	0.66
Using in-school expertise	16	2.03	0.64	17	1.94	0.56	16	1.75	0.68	18	1.89	0.63
Response teacher feedback	10	2.20	0.79	9	2.11	0.78	12	2.08	0.51	9	1.56	0.73
Identifying assessment data	17	2.59	0.92	18	2.42	1.02	18	2.69	0.96	18	2.17	1.03
Using assessment data	12	2.63	0.48	9	1.94	0.81	17	2.56	0.50	18	2.36	0.64
Average score	18	2.16	0.40	18	1.99	0.32	18	2.32	0.31	18	2.07	0.30

Scores potentially range 1-4 on each item

needed to be developed. In addition there was confusion about purposes that could be addressed with a goal setting activity directed at developing a framework for team and individual consulting.

Study Year 2 (1989-1990)

Sample/Procedure

Ten boards in eastern and central Ontario ranging in size from 6,800 to 52,600 students agreed to participate by sending a team consisting of one experienced consultant who would act as team leader and two to four inexperienced consultants. There were 35 volunteers in the program; all were viewed as competent by their supervisors. The inservice program was delivered in rotating sites with different boards serving as hosts for each of the five sessions (three full days, two half days).

The consultant sample was relatively young: most were aged 36-40 and only four were over 50. Two thirds were female and half the total sample had interrupted their careers, in most cases for family reasons. Most had not been in their current role for very long: only three had more than three years experience in their current position and almost half had none. When previous consulting experience was added to current tenure, the mean was 2.66 years as a consultant. Only two had consultancy experience outside their own board. A majority had responsibility for helping both elementary and secondary teachers; the remainder worked solely in the elementary panel, with more in the primary than in the other divisions. Most were very experienced teachers (17 years on average); five had served as principal or vice-principal prior to becoming a consultant. Most had served on various board curriculum committees although a fifth had not. Formal qualifications were biased toward professional rather than academic training. Virtually all had obtained at least one subject specialist certificate and most reported additional practical qualifications. Five had not completed an undergraduate degree and only three had acquired a master's. One had finished a doctorate.

The Year 2 sessions, delivered from late October 1989 to late April 1990, were similar to those in Study Year 1 with refinements in materials and procedures in response to feedback from Year 1 participants. A main goal of the revisions was to provide better focus. For example, in Year 1 both of the selected dimensions for a given pair of case studies were highlighted in the questions following each case description; in Year 2 the questions for one case highlighted one dimension, and the second case highlighted the other dimension. Another way of achieving better focus was by addressing fewer aspects of a profile dimension in a single session, providing participants with a clearer sense of inservice parameters, and designing new activities to help consultants set personal learning goals. The increased focus made it possible to provide additional practice activities for each dimension without extending the time allotted to each session.

Each board team had a formally designated team leader who was assigned the task of guiding the team through the inservice activities and stimulating reflection within the group. The researchers observed the team's deliberations but did not intervene except to respond to explicit requests for procedural direction. Separate sessions with team leaders were used to review progress, to plan subsequent activities, and to provide additional information. With

respect to the latter, researcher meetings with team leaders were used to introduce and discuss the two Profile dimensions that Study Year 1 data indicated were poorly grasped by the participants. These dimensions concerned with individualizing and responding to feedback, respectively, within consultation, were discussed in the context of helping team members develop an awareness of when on-the-spot adjustments are demanded by the consulting situation.

Pre and post interview data were collected[3] from the participating consultants using the same instrument as in Year 1. The pretest interviews were conducted in September 1989; the post interviews were administered in May 1990. All interviews were audiotaped. Two trained research assistants coded the tapes, assigning Consultant Profile levels to each subject on the basis of the examples of individual and group consulting that were provided. Any disagreements between coders were resolved through discussion. In addition, each consultant was asked to nominate a teacher familiar with his or her work. The teacher was asked to rate the performance of that consultant on each of the dimensions of the profile, using a five point scale that included "don't know." These teacher ratings were summed and correlated with the sum of the profile placements from the interview.

When all interviews were transcribed and coded, pre and post means and standard deviations were generated. T-tests were used to assess the significance of pre/post changes. The data from Year 2 were combined with the data from Year 1 in an analysis of covariance where the dependent variable was post interview score, the independent variable was treatment condition (Year 1 control, Year 1 treatment, and Year 2 treatment) and pretest score was the covariate.

In June a confidential report was sent to each participant describing the inservice team's perception of the extent to which the individual's strategies had changed during the program. Specific evidence for these perceptions was provided in the form of excerpts from the pre and post interviews. Subjects were invited to add missing data and/or correct interpretation of this data; no additional feedback was received.

Results

There were a large number of uninterpretable reactions to the interview questions on "response to teacher feedback during delivery." Given the number of cases for which one or both scores was missing, the item was deleted. The remaining 10 items were summed to produce a total score. All items correlated positively with the sum of the remaining items; the alpha reliability on pretest scores (n=28) was .58; on the post (n=33) it was .79. The sum of the items also correlated positively with the sum of the teacher ratings (r=.67) but the correlation was based on only 15 cases because of missing data.

Table 2 displays the means and standard deviations for each dimension and the total score. The table shows that Consultant Profile scores increased from pre to post; the differences were significant overall and on eight of the 10 dimensions.

Comparisons were also made between the scores of the Year 2 treatment condition and the Year 1 treatment and control groups. The dependent variable was the posttest average of nine profile interview items that were com-

Table 2
Means and Standard Deviations for 1990 Study Outcome Measures

Profile Dimension	Pretest N	Pretest $\bar{X}$	Pretest SD	Posttest N	Posttest $\bar{X}$	Posttest SD
Identifying consultation recipients	35	1.69	0.53	35	2.17	0.84***
Recruiting consultation team	35	1.96	0.85	35	2.47	0.74**
Determining needs	35	2.73	0.81	35	2.83	0.79
Individualizing within consultation	32	1.63	0.99	33	2.20	1.19**
Practice within sessions	33	2.46	0.93	35	3.20	0.69***
Practice between sessions	33	1.95	0.71	35	2.39	0.69**
Using in-school expertise	31	2.00	0.82	35	2.34	0.85
Identifying assessment data	35	1.89	0.83	35	2.54	0.93***
Using assessment data	32	2.28	0.70	35	2.76	0.65**
Reflecting within the role	35	1.83	0.38	35	2.03	0.45*
Total	28	20.91	3.56	33	24.97	4.77***

Scores potentially range 1-4 on each item and 4-40 overall.
*p<.05. **p<.01. ***p<.001.

mon to the three data sets. (The "reflecting within the role" dimension was not included because it was missing from Year 1 data; "responding to teacher feedback during delivery" was omitted because it was missing from Year 2 data.) The analysis of covariance showed that posttest scores were strongly influenced by pretest performance [$F(1,68)=23.25$, $p<.001$] and to a lesser degree by treatment condition [$F(2,68)=5.35$, $p<.007$]. The Newman-Keuls procedure for the multiple range test indicated that the Year 2 treatment group learned significantly more than the Year 1 control and the Year 1 treatment. The revision of the materials and procedure used to implement the treatment in the second administration appeared to have a positive effect. In this analysis the differences between the Year 1 treatment and Year 1 control did not reach significance, in contrast with the small but significant finding reported in Table 2, because the dependent measure was based on nine rather than 10 items.

As in the Year 1 study, observations made during the inservice sessions were used to refine the program further. For example, it was observed that labels (e.g., "the empty request", "the hostile audience") that had been used as case study titles should be deleted because they were a distraction to some participants.[4]

Discussion

The primary purpose of the Year 2 study was to find out if revisions to the inservice materials and procedures had a positive effect on performance. Comparison with the 1989 treatment and control groups indicated that the 1990 treatment was superior. The revision of inservice materials and procedures had a positive effect, most likely by increasing the reflectiveness of the participants while working through the professional development tasks. This

argument, along with audiotape evidence on consultant interactions during the inservice, is expanded in Ross and Regan (1991b).

General Discussion

The findings from this research will be of primary interest to those concerned with the enhancement of the consultant role. The findings suggest that it is possible to improve the quality of consultants' strategies through a collaborative program of input, reflection, and action based on the inservice principles described above. But the findings will also be of interest to those who are experimenting with the use of case studies to stimulate professional growth. Several collections of teacher case studies have been recently published (Greenwood & Parkay, 1989; Kowalski, Weaver, & Henson, 1990; Shulman & Colbert, 1987, 1988; Silverman, Welty, & Lyon, 1991) for use in preservice and inservice programs. There have been few evaluations of the case study approach to teacher development, although some evidence of positive impact has been provided by Welty, Silverman, and Lyon (1991), and to a lesser extent by Shulman and Colbert (1989). The research reported here corroborates the claim that reading and discussing the experiences of other professionals can make a positive contribution to professional growth. The findings, particularly the differences in the outcomes of the 1989 and 1990 versions of the program, also suggest that certain conditions play a critical role in the effective use of case studies.

1. The problems elaborated in cases must be so familiar to inservice participants that they are able to spontaneously place themselves in the setting. Without such familiarity the connections between one's own experience and that of others as a foundation for new learning is unlikely to be made. In this program the cases were generated from reports of consultants working in similar settings to those of the inservice participants, and specific discussion probes were provided to help them make connections between the case, their past actions, and future intentions.
2. Each case should be more than a report of an event: it should represent an instance of a larger class and illustrate a particular principle of practice (Shulman, 1986). McAninch (1991) criticized existing collections of teacher case studies for failing to make explicit linkages between case content and theories of what teachers should know and do. In the consultant inservice the cases were written to focus attention on specific dimensions of consultant strategies that distinguished relative experts from novices.
3. Learners need to be motivated by the case discussion to change their practice. In the 1989 study some groups questioned whether strategies described in the higher levels of the Profile were superior to their current methods. In the 1990 version the discussion probes were revised so that groups did a card sort in which they rank ordered alternative ways of dealing with the problem outlined in the case. When the Profile dimensions were distributed, consultants invariably found that their rank ordering of the approaches conformed to the sequence in the Profile. This activity enhanced the credibility of the Profile, displaced their commitment to existing routines, and motivated change in professional practice.
4. The case studies should provide an achievable challenge. The cases used in the consultant inservice were not worst case scenarios. Rather, they sug-

gested how in a given situation an inexperienced but conscientious consultant without a solid repertoire of strategies might respond to a significant problem. Consultants had to make use of their own experience and professional knowledge, as well as data from the research literature, to come up with better solutions for case study problems.

5. The length of the case studies may influence inservice learning. At times participants felt not enough information was presented in the cases and they rationalized the behavior of the consultant described in the case by suggesting that there may have been unknown factors preventing the protagonist from acting differently. At other times case study details cued discussion of some aspect of consulting irrelevant to the task at hand (e.g., discussion of attitude/behavior of a particular principal). The case studies used in both years of the consultant inservice were considerably shorter (about 200 words prior to discussion probes) than most teacher case studies that have been reported.

Our findings, especially in Year 2 of the study, lead us to conclude that the systematic treatment of new learnings initiated in the case studies in each session, and carefully elaborated in the planning, practice, and subsequent reflections on practice, offer some explanation of the success of the inservice treatment. Certainly other elements might affect achievement such as the duration of the program, the training of group leaders, the goals and objectives of sessions, grouping strategies, and, as already suggested, procedures for feedback and reflection. The relative contribution of each factor in the model to the final outcome was not addressed in the research reported here and further studies systematically varying each design element would be appropriate.

Notes

1. Funding for this research was provided by grants from the Ontario Ministry of Education to the Ontario Institute for Studies in Education. The views expressed in the article do not necessarily reflect the views of the Ministry. The authors thank Anne Hogaboam-Gray who contributed to the data collection and Peter Gray who contributed to the data analysis. The authors would also like to express their appreciation to the consultants who participated in the research. A version of the article was presented at the annual meeting of the Canadian Association for Curriculum Studies, Kingston, Ontario in June 1991.

2. Interview data were supplemented with additional information that are not reported here. These included (a) records of action plans and reports of implementation experiences generated during inservice activities; (b) field notes produced by researchers during the sessions; (c) job descriptions, workshop outlines and the like provided by consultants; and (d) audiotapes of deliberations among consultants when engaged in case study analysis, action planning and feedback tasks.
3. Data were also collected on antecedent conditions that affected consultant learning during the inservice. These are reported in Ross and Regan (1991a).
4. Other differences between Year 1 and Year 2 implementations of the program, are reported in Ross and Regan (1991b).

References

Bereiter, C., & Scardamalia, M. (1986). Educational relevance of the study of expertise. *Interchange, 17*(2), 10-19.

Berliner, D. (1986). In pursuit of the expert pedagogue. *Educational Researcher, 15*(7), 5-13.

Broskowski, A., Khajavi, F., & Mehryar, A. (1973). An evaluation of an inservice training seminar on community consultation and education. *Journal of Community Psychology, 1*(2), 174-176.

Clark, C. (1988). Asking the right questions about teacher preparation: Contributions of research on teacher thinking. *Educational Researcher, 17*(2), 5-12.

Coffing, R., & Hutchison, T. (1974). *Needs assessment methodology: A prescriptive set of rules and procedures for identifying, defining, and measuring needs.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Curtis, M., & Watson, K. (1980). Changes in consultee problem clarification skills following consultation. *Journal of School Psychology, 18,* 210-221.

Firestone, W., & Corbett, H. (1981). Schools versus linking agents as contributors to the change process. *Educational Evaluation and Policy Analysis, 3*(2), 5-17.

Froehle, T. (1978). Systematic training for consultants through competency-based education. *Personnel and Guidance Journal, 56,* 436-441.

Glass, G., McGraw, B., & Smith, M. (1981). *Meta-analysis in social research.* Beverley Hills, CA: Sage.

Gliessman, D.H., Grillo, D.M., & Archer, A.C. (1989, March). *Changes in teacher problem solving: Two studies.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Goodwin, D., & Coates, T. (1973). Increasing teacher effectiveness through social systems change: Training school psychologists as change agents. *California Journal of Educational Research, 25*(3), 147-156.

Greenwood, G., & Parkay, F. (1989). *Case studies for teacher decision making.* New York: Random House.

Griffin, G., & Barnes, S. (1986). Using research findings to change school and classroom practices: Results of an experimental study. *American Educational Research Journal, 23,* 572-586.

Griffin, W. (1978). Educational needs: Definition, assessment and utilization. *School Review, 86,* 382-394.

Kagan, D. M. (1990). Ways of evaluating teacher cognition: Inferences concerning the Goldilocks jungle. *Review of Educational Research, 60,* 419-469.

Kowalski, T., Weaver, R., & Henson, K. (1990). *Case studies on teaching.* New York: Longman.

Leithwood, K., & Stager, M. (1987). *Components of expertise: "Artistry" in principals' problem solving.* Paper presented at the Canadian Society for the Study of Education, Guelph.

Loadman, W., & Mahan, J. (1972). The external consultant and curriculum change strategies. *Theory into Practice, 11*(5), 329-339.

Loucks, S. (1982). External and local facilitators. In S. Loucks (Ed.), *People, policies and practices: Examining the chain of school improvement. Vol. 4: A study of dissemination efforts supporting improvement* (pp. 43-92). Andover, MA: The Network.

Louis, K. (1981). External agents and knowledge utilization: Dimensions for analysis and action. In R. Lehming & M. Kane (Eds.), *Improving schools: Using what we have* (pp. 168-211). Beverly Hills, CA: Sage.

Louis, K., & Kell, D. (1981). *The human factor in dissemination: Field agent roles in the organization context. Linking R&D with schools.* Cambridge, MA: ABT. (ERIC Document Reproduction Service No. ED 207 266)

Louis, K., & Sieber, S. (1979). *Bureaucracy and the dispersed organization: The educational extension experiment.* Norwood, NJ: Ablex.

McAninch, A. (1991). Casebooks for teacher education: The latest fad or lasting contribution. *Journal of Curriculum Studies, 23*(4), 345-355.

Medway, F.J. (1979). How effective is school consultation? A review of recent research. *Journal of School Psychology, 17,* 275-282.

Miles, M., Fullan, M., & Taylor, G. (1978). *Organizational development in schools: The state of the art. Vol. III: OD consultants/OD programs in school districts.* New York: Center for Policy Research.

Miles, M., Saxl, E., & Lieberman, A. (1988). What skills do educational "change agents" need? An empirical view. *Curriculum Inquiry, 18*(2), 157-193.

Ontario, Government of. (1984). *Regulation 262. Revised regulations of Ontario, 1980, as amended to O. Reg. 761/82 under Education Act (Elementary and secondary schools and schools for trainable retarded pupils—general).* Toronto: Queen's Printer.

Padfield, C. (1981). Why school boards should have consultant development programs. *Education Canada, 21*(2), 38-41.

Regan, E., & Ross, J.A. (1991). *Professional development activities for the reflective consultant: Techniques for learning how to help schools.* Toronto: OISE Publications.

Ross, J.A., & Regan, E. (1990). Self-reported strategies of experienced and inexperienced consultants: Exploring differences. *Alberta Journal of Educational Research, 36,* 157-180.

Ross, J.A., & Regan, E. (1991a). *Antecedent conditions affecting learning in an in-service program for consultants.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Ross, J.A., & Regan, E. (1991b). *Sharing professional experience: Talk and achievement in adult cooperative learning groups.* Mimeo. Peterborough, ON: OISE Trent Valley Centre.

Runckell, P., & Bell, W. (1976). Some conditions affecting a school's readiness to profit from OD training. *Education and Urban Society, 8,* 127-144.

Schön, D. (1983). *The reflective practitioner: How professionals think in action.* New York: Basic.

Schön, D. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.* San Francisco, CA: Jossey-Bass.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Shulman, S., & Colbert, J. (Eds.). (1987). *Mentor teacher casebook.* San Francisco, CA: Far West Laboratory for Educational Research and Development.

Shulman, S., & Colbert, J. (Eds.). (1988). *Intern teacher casebook.* San Francisco, CA: Far West Laboratory for Educational Research and Development.

Shulman, S., & Colbert, J. (1989). Cases as catalysts for cases: Inducing reflection in teacher education. *Action in Teacher Education, 11*(1), 44-52.

Silverman, R., Welty, W., & Lyon, S. (1991). *Case studies for teacher problem solving.* New York: McGraw-Hill.

Tyler, M., & Fine, M. (1974). The effects of limited and intensive school psychologist-teacher consultation. *Journal of School Psychology, 12*(1), 8-16.

Welty, W., Silverman, R., & Lyon, S. (1991). *Student outcomes from teaching from cases.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 3, September 1992, 205-218*

Marziya Yasmin

and

Baha Abu-Laban
University of Alberta

Ethnicity and Occupational Inequality: A Reconsideration

There is disagreement in the current literature on social stratification concerning whether ethnicity has an impact on socioeconomic status. This article addresses this issue by introducing a new conceptual-methodological perspective. Using data from the Public Use Sample Tapes of the 1981 Canadian Census, the study employs multiple indicators of ethnicity, incorporates a broad set of social background variables into the research design and examines both inter- and intra-ethnic occupational inequality. The results indicate that social background variables, in particular gender and education, are more important than ethnicity in explaining occupational inequality. However, the results also suggest the continuing relevance of ethnicity in explaining some aspects of inequality. The research and practical implications of these results are discussed.

Il y a un désaccord dans la littérature actuelle de la stratification sociale concernant l'effet de l'appartenance ethnique sur le statut socio-économique. Cet article traite de cette question en introduisant une nouvelle perspective conceptuelle et méthodologique. Se servant de données provenant des Public Use Sample Tapes *du recensement canadien de 1981, l'étude utilise des indicateurs multiples de l'appartenance ethnique, incorpore un large éventail de variables d'ordre social dans le plan expérimental et examine aussi bien l'inégalité inter-ethnique que l'inégalité intra-ethnique dans l'emploi. Les résultats indiquent que les variables d'ordre social, en particulier le sexe et l'éducation, sont plus déterminantes que l'appartenance ethnique dans l'explication de l'inégalité dans l'emploi. Cependant, les résultats suggèrent aussi la pertinence de l'appartenance ethnique dans l'explication de quelques aspects de l'inégalité. La recherche et les implications de ces résultats sont présentées.*

Introduction

The issue of ethnic inequality has been a major research focus in Canada in recent years. Much of the credit for this goes to Porter (1965) whose seminal work has stimulated considerable interest and set the stage for an ongoing debate in this area. Porter maintains that ethnicity has played an important role in class formation in Canada:

> Immigrants from many different cultures have had to fit into a social class structure which, in Canada, is a reflection of the economic position of the many ethnic groups, both Canadian-born and immigrant, who make up the population. In this selecting and sorting out of migrants of different ethnic back-

Marziya Yasmin is a research associate in the Department of Sociology.
Baha Abu-Laban is a professor of sociology and Associate Dean (Research and Graduate Students) in the Faculty of Arts.

> grounds into various occupations and so into the class system, a number of factors have operated in varying intensities at different times in Canada's history. (p. 60)

A main theme of Porter's work concerns the persistence over time of a general pattern of vertical arrangement of ethnic groups. According to him:

> Germans, Scandinavians, and Dutch are nearest to the British in their occupational levels ... Italians, Polish, Ukrainians, and groups from southeast Europe are still at the lower end of the occupational spectrum ... The French could be placed somewhere between these last and the groups from northern Europe. (p. 90)

Porter (1965, p. 68) cogently argues that the prevailing racial attitudes in Canada operated in such a way as to funnel the less preferred immigrants (e.g., East Europeans) into a lower occupational or "entrance status"; and to restrict their mobility and that of their descendants. He further argues that in both the industrial and agricultural sectors, ethnic differences have been important in the expansion of the bottom layer of the stratification system. Although there may be local and regional variations in the relationship between ethnic affiliation and social class, the general pattern remains. Apart from everything else, Porter clearly saw rising tension between ascription and achievement as two distinct principles of resource allocation.

The substance of the current debate concerning Porter's thesis is whether ethnic differences do play such an important role in class formation. The works of Richmond (1967a; 1967b), Blishen (1970), Li (1978), Forcese (1980), Richmond and Kalbach (1980) and Lautard and Loree (1984) tend to support Porter's thesis. In contrast, however, the works of Pineo (1976), Darroch (1980) and Reitz (1980) have demonstrated a weak influence of ethnicity on socioeconomic status. More recent research evidence reported by Satzewich and Li (1987) shows that income disparity among ethnic groups has tended to persist despite an appreciable reduction in occupational disparity.

A more challenging position has been taken by Li (1988) in his study of ethnic inequality, where he argues that inequality can be better explained by class relations than by ethnicity. Li acknowledges that both class and ethnicity produce inequality in the capitalist system, but

> they tend to operate differently. The source of class inequality.... stems from the unequal relationships to production. Ethnic inequality, however, surfaces in the capitalist labour market. In other words, class-based inequality is fundamental in capitalist production, whereas ethnic and racial divisions are not the direct result of relations of production. (pp. 51-52)

Thus a substantial portion of income disparity may be attributed more to class than to ethnicity (Li, 1988, p. 138).

Although Li's (1988) study presents an interesting articulation of class and ethnicity, it nevertheless raises important questions regarding the conceptual and methodological issues involved in this area of study. These issues draw attention to the need for the development of a new research strategy.

Using a strategy that involves a less limiting research procedure, this article addresses three major areas of concern in the study of ethnicity and occupational inequality in Canada. First, it seeks to reassess Porter's thesis that ethnic

affiliation is a correlative factor in the assignment of occupational roles and thus in the class system. Second, it seeks to determine the extent to which ethnicity, in comparison with other factors, helps to explain occupational disparity. Finally, in order to probe the dynamics of inequality a little more fully, the article seeks to identify correlative factors in occupational inequality within ethnic groups (i.e., intra-ethnic occupational inequality). The overall results from this study may shed some light on whether the above noted contradictions in the literature may be attributed to methodological differences or to a genuine change in class dynamics in Canadian society.

Critical Reflections

Li (1988) correctly notes that "one of the problems surrounding the debate of the vertical mosaic concerns the different ways ethnicity and class are measured in each study" (p. 61). With reference to ethnicity, although it is a compound concept made up of both subjective and objective elements, it has often been defined in terms of only one element, be it subjective or objective. An example of a subjective definition is *ethnic identification*, whereby the individual expresses a positive personal attitude toward a given group. On the other hand, an objective definition involves *classification* by others on the basis of physical (e.g., skin color) and/or sociocultural characteristics (such as language or religion). Moreover, some definitions of ethnicity continue to be characterized by a high degree of ambiguity. For example, in the 1981 Canadian Census, which is a major source of information for social scientists, ethnic origin is determined by the following question: "To which ethnic or cultural group did you or your ancestors belong on first coming to this continent?" The responses to this question are difficult to interpret to say the least, particularly since the respondents were allowed to indicate single or multiple origins, with no instruction given on whether to trace ethnic origin through the father's or mother's side.

The works of some of the earlier researchers such as Porter (1965), Darroch (1980), Lautard and Loree (1984) who followed the Canadian Census definition of ethnic origin in their studies, raise much concern about the validity of conceptualizing ethnic origin. Other researchers like Blishen (1970), Richmond (1967a; 1973) and Boyd (1985a) used country of birth or national origin for examining inequality. In this connection, it has been pointed out by Pineo and Porter (1985) that ethnicity loses much of its meaning among native-born Canadians with subsequent generations, which restricts the use of this variable to the foreign-born members only. Ethnic language was another indicator that some studies employed for examining ethnic inequality as evidenced by the works of Tepperman (1975), Boulet and Veltman (1981) and McRoberts (1985). Again, like place of birth, language is likely to be of significance mainly to foreign-born respondents.[1]

Although each of these indicators is more or less limited in defining ethnicity, it is reasonable to assume that the use of all three indicators, that is, ethnic origin, place of birth, and language simultaneously, may provide a more comprehensive definition of ethnicity. The value of using multiple indicators of ethnicity is shown in the works of Goldlust and Richmond (1974) and Ornstein (1981), even though these studies are restricted to small samples. More recently, Li (1988) employed ethnic origin and nativity for examining

inequality, but the thrust of his analysis appears to be more on ethnic origin than on nativity. The present research employs all three of the indicators mentioned above and includes a large national sample.

With reference to the measurement of class, there are major differences among the earlier studies. Few of these studies have conceptualized class in Marxian terms, that is, in terms of relations of production. For this reason, Li's (1988) study is an important contribution in this regard. Li's analysis of class is based on employment status (as defined in the 1981 Canadian Census). The three categories of employment status included in his study are "self-employed with paid help," "self-employed without paid help" and "employees or workers." The latter category, "employees or workers," has been further subdivided into "managers and supervisors," "professionals" and "workers." Li maintains that employees engaged in managerial or supervisory occupations have some control over money capital and considerable control over physical capital which refers to actual production and labor of others. However, a detailed examination of the managerial category indicates wide occupational differences ranging from production management to post office management. Clearly the degree to which employees in this category exercise control over capital and labor is extremely variable. This is a serious drawback of Li's classification scheme.

Moreover, Li's (1988) attempt to examine the impact of social class on income is tautological in the sense that one of the underlying components in conceptualizing class position is ownership of property, which is often indicated by income.[2]

One of the most widely accepted ways of measuring class position involves the use of such indicators as occupation, income, and education, or often a combination of these variables commonly referred to as socioeconomic status. Although this method is gradational in nature and hence not grounded in a theory of class as a relational concept, the resulting hierarchical categories would most likely reflect significant economic and behavioral differences among the people (Porter 1965, p. 10). The present study follows the gradational mode of class analysis, but for technical reasons it uses occupational rather than socioeconomic status.[3]

Turning now to other issues, a further limitation observed in earlier studies concerns the bivariate analyses of ethnicity and social class. Such an approach tends to overlook the effect of ethnicity on social class while controlling for the effect of additional variables. To overcome this, proponents of the status attainment model (Blau & Duncan, 1967; Cuneo & Curtis, 1975; Duncan & Duncan, 1968; Li, 1978; and Boyd et al., 1981) sought to examine the impact of ethnicity on socioeconomic status by including certain social origin and achievement oriented variables. In this context, a number of researchers have elicited the importance of level of education in explaining ethnic inequality. For example, Li (1978) demonstrates that occupational differences among male immigrants are determined in part by differences in education. Other researchers such as Ornstein (1981) have demonstrated that education plays a more important role than ethnicity in explaining occupational disparity. The present study deals with this aspect by examining first the role of ethnicity in explaining differences in level of education and, second, the role of ethnicity in explaining occupational inequality by controlling the effects of education.

A systematic review of earlier studies also reveals that they have tended to neglect the inclusion of demographic variables. To further our understanding of ethnic inequality, it is important to widen our framework of inquiry and include important demographic variables. Studies utilizing this approach have been sporadic and often confined to a single demographic factor. For example, some researchers have examined the degree to which inequality in socioeconomic status can be explained by gender (Abella, 1984; Boyd, 1985b; 1990; and Fox & Fox, 1986); or by age (McRoberts et al., 1976; and Pineo & Porter, 1985); or by family size (Kenneth, 1970; Pineo, 1985; Trovato & Grindstaff, 1986); or mobility status (Blau & Duncan, 1967; Sandefur & Scott, 1981; and Sharma, 1980). In this study an attempt is made to introduce all these demographic variables along with ethnic variables to examine the net effect of ethnicity on socioeconomic status.

The mode of inquiry followed by researchers of ethnic inequality also warrants some attention. Two main perspectives were followed in this regard. First, some researchers followed an exploratory mode of inquiry and examined the extent to which ethnicity had an influence on social class (Blishen, 1970; Darroch, 1980; Lautard & Loree, 1984; Ornstein, 1981; Porter, 1965). Other researchers approached their investigation by assuming prevalence of disparity in socioeconomic status among ethnic groups and thus proceeded to explain intra-ethnic inequality (Cuneo & Curtis, 1975; Boyd et al., 1981; Boyd, 1985a). The first mode of inquiry can be regarded as an important step toward understanding the role of ethnicity in explaining the dynamics of class inequality, whereas the second approach can be viewed as critical for analyzing class and inequality within various groups. Thus the two approaches can be regarded as complementary; jointly they provide a more comprehensive understanding of the interplay between ethnicity and social class. The present study, therefore, includes both of these research foci.

Admittedly the above noted problems are not exhaustive of the conceptual and methodological issues confronting this research area. However, we are in a position to address them in the context of the present research design. Specifically, we will examine (a) the influence of ethnicity, defined in terms of three indicators, on occupational status; (b) the net influence of ethnicity on occupational status, while incorporating a set of demographic variables in a broader model of ethnic variables, and (c) intra-ethnic occupational inequality.

Research Procedure

Data for this study were obtained from the *Public Use Sample Tapes* of the 1981 Census of Canada, representing two percent of the original sample and including individuals in the age range of 15 to 65 years (N=333,576). This age range represents the work force in the Canadian Census. The variables included in the study are discussed below, but it should be noted that the definitions of all these variables were derived from the Dictionary of the 1981 Census of Canada.

Ethnicity. As indicated earlier, an important feature of this study is that ethnicity is defined in terms of three indicators, namely, ethnic origin, place of birth, and use of home language. While all three of these indicators have been used in the studies reviewed above, the tendency has been to define ethnicity in terms of one of them and not all three together. Given the compound nature

of the concept, we believe that the use of multiple indicators in this instance is methodologically less limiting than the use of only one. The three indicators used in the study are explained briefly below.

Ethnic origin refers to the ethnic group to which the respondent or the respondent's ancestors belonged on first coming to Canada. Six ethnic groupings, primarily based on their broad geographic origin and period of arrival, have been identified. These are the British, French, North-West Europeans, South-East Europeans, non-Europeans, and a final grouping made up of individuals with multiple origins.

Place of birth, also referred to as nativity, is the second indicator of ethnicity. This was entered as a dichotomy of foreign-born and native-born.

The Canadian Census provides information on both mother tongue and home language. Whereas mother tongue refers to the language first learned in childhood and still understood by the respondent, home language refers to the language actually spoken at home by the respondent at the time of the census. For the purpose of this study, home language was regarded as the more valuable indicator of the two. Home language, the third indicator of ethnicity, was entered as a trichotomy, including use of English, use of French, and use of other languages at home.

Occupation. The preceding review of the literature shows that ethnic inequality has been variously studied in terms of class, socioeconomic status, or occupational inequality. As discussed above, the measurement of class from census data seems to present serious problems. Also, although socioeconomic status is a widely accepted measure, there were technical problems in using it in this study[3]. Thus it was concluded that for present purposes occupation was the most useful measure for examining ethnic inequality. This conclusion is not without merit since it is based on the widely held assumption that a person's occupation very often mirrors his or her status or position in society and forms an important criterion for stratification. Moreover, occupation tends to reflect both education and income.

Occupation was entered by using 16 categories of the Canadian Classification and Dictionary of Occupations. These were: Managerial, Natural Sciences, Social Sciences, Teaching, Medicine and Health, Artistic/Literary, Clerical, Sales, Service, Farming, Other Primary Occupations, Processing Occupations, Machining and Repair, Construction Trades, Transport/Equipment, and Other Occupations. After testing these categories in relation to income and education, we collapsed them into three manageable categories representing high, medium, and low occupational status. Of the 16 categories listed above, the first six were combined to define high occupational status; the next three (Clerical, Sales, Service) were used to identify medium occupational status; and the remaining seven categories were combined to identify low occupational status. These occupational status categories were found to resemble those identified by Pineo, Porter and McRoberts (1977, p. 95).

Demographic (social background) variables. Another important feature of this study is the attempt to examine the relationship between ethnicity and occupational status while controlling for the effects of certain demographic or social background variables. The control variables incorporated into the design of the present study are those indicated by our critical review of the literature, namely: age, gender, family size, mobility status, and level of educa-

tion. Leaving age and gender aside, the census definitions of the remaining variables are as follows.

Family size refers to the number of persons in a family unit (as defined by the 1981 Canadian Census). Ten categories of this variable have been entered, ranging from one to 10 members in the family.

Mobility status has been dichotomized between movers and nonmovers, with movers being individuals who did not live in the present residence five years ago.

Level of education refers to the highest level of schooling attained and has been entered as an ordinal variable with 11 categories.

Multiple regression analysis. The procedure of multiple regression was employed for analyzing the data because this procedure allows the researcher to examine the effect of the independent variables on the dependent variable while controlling the effects of other variables in the model. This is a powerful statistical procedure well suited for analyzing the relationship between ethnicity and occupational inequality as it helps to overcome the limitations of the bivariate analyses of ethnic inequality that have been followed in much of the literature.

The analysis is based on standardized (*B*) rather than unstandardized (b) regression coefficients. Despite the relative instability of standardized coefficients across samples, they have been used in the analysis because they provide a more appropriate means of interpreting the effects of variables using different scales. As Pedhazur (1982) observes:

> Unlike the b's, *B*'s are scale-free and therefore can be compared across different variables. It is this property of the *B*'s that appeals to a great many researchers, who use their relative magnitude as an indication of the relative importance of the variables with which they are associated. (p. 247)

To sum up, the research strategy followed in this study is unique and arguably less limiting because it combines the following features: use of a large national sample, use of multiple indicators of ethnicity, incorporation of numerous background variables into the research design, examination of inter- and intra-ethnic occupational inequality, and use of multiple regression.

The Findings

In the light of earlier viewpoints regarding the two modes of inquiry into ethnic inequality, the analysis was carried out in two stages. The first stage treated ethnicity as a determinant of occupation and education and examined the extent to which ethnic and demographic variables explained variation in occupational status and level of education. In the second stage, factors explaining intra-ethnic occupational inequality were examined.

Considering the first stage of analysis, Table 1 shows two different models that were employed for explaining the influence of ethnic and demographic variables on occupational status and level of education respectively.[4] The results of model one are presented in columns 2 and 3. The table shows that this model explains 15% of the variance in occupational status. On the whole, the variables representing ethnicity (i.e., ethnic origin, nativity and home language) appear to be weak in accounting for variation in occupational status, though use of English as a home language appears to have the strongest

Table 1
Occupational and Educational Inequality by Selected Ethnic and Demographic Variables

	Occupation		*Education*	
	b	*B*	b	*B*
Ethnicity Variables				
Ethnic Origin				
British	−.011	−.009+	−.036	−.056*
French	.005	.003+	−.072	−.102*
NW European	.016	.007**	.019	−.017*
SE European	.058	.027*	−.071	−.066*
Multiple Origins	−.032	−.014*	−.013	−.011+
Nativity	.062	.019*	.161	.098*
Home Language				
English	.065	.049*	.118	.174*
Education	.420	.212*	—	—
Demographic Variables				
Gender	.684	.273*	.063	.049*
Age	−.004	−.055*	−.003	−.069*
Family Size	−.060	−.065*	−.032	−.070*
Mobility	.119	.047*	.121	.095*
Constant	.703		1.292	
Multiple R	.392		.204	
R^2	.154		.041	
Adj R^2	.154		.041	
Std. Error	1.150		.620	
DF	12/333563		11/333563	
F	2292.607		592.086	
N	333576		333575	

+Statistically not significant.
*Statistically significant at .01 level.
**Statistically significant at .05 level.

influence (*B.049*) on occupational status. This means that with an increase in the proportion of respondents who use English as home language, occupational status tends to rise.

Among the demographic variables, gender is the most influential variable in explaining variation in occupational status (*B.273*), which in this case means that with an increase in the proportion of male members there is a likely rise in occupational status. Level of education is the next most important variable influencing occupational status (*B.212*), indicating that with an increase in the proportion of those with a high level of educational attainment, there is a concomitant rise in occupational status. The influence of family size and age on occupational status (*B–.065* and *B–.055* respectively) is relatively weak. The negative coefficients of these variables indicate that with an increase in size of families and with age, occupational status tends to decline. Mobility status has the weakest influence on occupational status, with a beta coefficient of .047. It is worth noting, however, that the positive association between mobility and occupational status supports earlier observations by Sharma (1980) and San-

defur and Scott (1981) and possibly reflects the expectation of occupational reward for those who are geographically more mobile.

A somewhat different picture emerges when the influence of ethnic and demographic variables on level of education is examined in model two. It is apparent from column 4 in Table 1 that only 4% of the variation in level of education is explained by this model (compared with 15% of the variation in occupational status explained by model one).

Although the same ethnic variables are used in models one and two, it is clear from Table 1 that these variables have a relatively stronger influence on level of education (model two) than on occupational status (model one, columns 2 and 3). With specific reference to model two, the most noteworthy finding is that use of English as home language has the strongest influence on level of education, with a beta coefficient of .174, which is higher than all the other beta coefficients in the model. This means that with an increase in the proportion of those who use English as home language, there is a likely rise in level of education. The influence of nativity on level of education is relatively weak (*B.098*). With reference to ethnic origin, the French and the South-East European groups show a strong negative association with level of education (*B–.102* and *B–.066* respectively). This finding shows some similarities with earlier observations by Ornstein (1981) to the effect that the French and the Italian ethnic groups had low scores on education.

Among the demographic variables, mobility status appears to be the most influential variable in explaining level of education (*B.095*), which suggests that with an increase in the proportion of members who are geographically mobile, there is a likely rise in level of education. This contrasts with a main finding of the previous model where gender and not geographical mobility was the most influential variable in accounting for variation in occupational status. This finding relating to model one does not necessarily mean that mobility status is of no relevance in explaining occupational inequality. Rather, it may suggest that those individuals (both male and female) who are geographically mobile are more likely to access educational opportunities, which in turn may positively influence their occupational achievement.

Finally, in model two, the influence of family size (*B-.070*) and age (*B-.069*) on educational attainment is relatively weak and that of gender is the weakest (*B.049*). These results probably reflect the democratization and hence greater accessibility of education in Canada today compared with the 1950s and early 1960s.

Turning now to intra-ethnic occupational inequality, Table 2 shows the effects of ethnic and demographic variables on occupational status.[5] Six regression models are presented, one for each ethnic group. The model representing the French ethnic group (column 3) explains 18% of intra-group variation in occupational status. For all the other models, the variance explained ranges from 13 to 14%.

Both the British and the North-West European ethnic groups show some similarity. For example, for both of these groups gender emerges as the most important variable in explaining within-group occupational inequality (*B.271* and *B.296* respectively), followed by level of education (*B.202* and *B.170* respectively). The variables next in order of importance are age, family size, and mobility status. Both use of English as home language and nativity have

relatively weak influence on occupational status. The French and the South-East European ethnic groups also demonstrate that gender has the strongest influence in explaining intra-group occupational inequality (*B.298* and *B.278* respectively), followed by level of education (*B.225, B.172* respectively). But for the French ethnic group, use of French as home language (*B–.074*) and family size (*B–.070*) appear to be the next most important variables in explaining internal differences in occupational status. The negative coefficient associated with use of French as home language indicates that with an increase in the proportion of French-Canadians who use French as home language, there is a likely decline, albeit small, in occupational status. For the South-East European group, nativity (*B.069*) and mobility status (*B.050*) are the next most influential variables in explaining intra-group variation in occupational status.

For the non-European ethnic group as well as those with multiple origins, level of education appears to be the most important variable in accounting for internal variations in occupational status (*B.235* and *B.237* respectively), followed by gender (*B.225* and *B.212* respectively). For the non-European ethnic group, use of English as home language (*B.084*) and nativity (*B.062*) are the next most influential variables affecting occupational status. But for members with multiple origins, family size (*B–.117*) is the most influential variable affecting occupational status, following level of education and gender.

A noteworthy feature that emerges from Table 2 is that for ethnic groups of British origin and those with multiple origins, the influence of home language

Table 2
Occupational Inequality by Nativity, Home Language, Gender, Age, Family Size, and Mobility Status of Six Ethnic Groupings

	British B	*French* B	*NW European* B	*SE European* B	*Non-European* B	*Mult. Origins* B
Ethnicity Variables						
Nativity	–.011	–.033	.004	.069	.062	–.014
Home Language						
English	.013	—	.050	.041	.084	.033
French	—	–.074	.016	–.012*	—	—
Education	.202	.225	.170	.172	.235	.237
Demographic Variables						
Gender	.217	.298	.296	.278	.225	.212
Age	–.090	–.021	–.085	–.041	–.004	–.036
Family Size	–.067	–.070	–.078	–.048	–.056	–.117
Mobility	.045	.062	.031	.050	.014	.035
Constant	.938	.249	1.250	.719	.632	1.128
Multiple R	.375	.426	.384	.362	.371	.361
R^2	.140	.182	.147	.131	.137	.130
Adj R^2	.140	.182	.147	.131	.137	.130
Std. Error	1.129	1.219	1.607	1.098	1.150	1.094
F	3041.729	2926.982	653.442	595.038	619.452	364.465
DF	7/130103	7/91938	8/30165	8/31370	7/27096	7/16984
N	130111	91946	30174	31379	21704	16992

Note: All the coefficients in this table are significant at the .01 level, with the exception of that denoted by an asterisk.

on intra-group occupational inequality is relatively weak. In contrast to this, the influence of home language on occupational status is somewhat stronger among ethnic groups of South-East European and non-European origin. It is interesting that these contrasting results mirror the degree of linguistic homogeneity among the different ethnic groups. For example, over 90% of British-origin Canadians as well as those with multiple origins use English as home language, whereas only 39.2% of South-East Europeans and 37.2% of those of non-European origin use English as home language (1981 Census of Canada). In the Canadian context, the striking differences in terms of home language within the South-East European and non-European ethnic groups are apt to be associated with differences in occupational status.

On the whole, the evidence from this study shows that for all ethnic groups demographic factors rather than ethnic factors are more important in explaining occupational disparity. Also, of all the demographic variables examined, gender and level of education have the strongest influence on occupational inequality. In fact a consistent pattern of strong influence of gender and level of education on occupational status emerges for all the models (except in model two where level of education is the dependent variable). This consistent pattern brings into focus the dynamic role of gender and level of education in the formation of occupational hierarchies in society at large and also across ethnic groups.

Conclusion

While this study reveals some interesting patterns, one would have to be cautious about overgeneralizing. For example, like most of the earlier studies, one of the limitations of the present study lies in using occupation as an indicator of (or substitute for) social class position. The question that remains unanswered is: If class measurement was properly grounded in the Marxian tradition of class theory, would ethnicity still play an important role in determining class inequality?

The inclusion of several demographic variables was an important aspect of this study. The results show that these variables, particularly gender, and also education tend to have a stronger impact on occupational status compared with ethnic variables. Significantly, these results in general are consistent with earlier studies that employed a multivariate rather than a bivariate research design. Judging from these results, we may be justified in concluding that the debate over Porter's thesis is to some degree a debate over method. With particular reference to gender, the evidence underlines the prevalence of a male dominated core labor market and directs attention to the need to investigate more thoroughly the dynamics of gender inequality in the labor market.

It is quite likely that since the publication of Porter's (1965) work the changing labor market conditions and the demand for recruiting immigrants with relatively high educational and occupational qualifications have mitigated the influence of ethnicity on occupational inequality. The start of this trend was evident to Porter (1965) when he declared that one of the reasons for the changes in relative social status "is the difference between ethnic groups in the occupational level of their immigrants" (pp. 86-88). For example, he underlined the exceedingly high representation of professionals (scientists and engineers) among Asian immigrants between 1951 and 1961. In

this context, it is significant to note that Canada's preference for highly qualified immigrants continued through the 1960s and 1970s. It is also quite likely that among native-born Canadians ethnic origin may carry a different meaning today than in the past. Pineo and Porter (1985), for example, in their study of current ethnic identity of Canadians, found that about 86% identified themselves as *Canadians*. These issues, notably the seemingly changing meaning of ethnicity, merit further investigation in the context of a comparative research design.

The fact that ethnic variables in general had a relatively weaker influence on occupational status than demographic variables is not to suggest the irrelevance of ethnicity in Canadian society. For one thing, use of English as home language was found to have a positive influence on level of education, which in turn played a critical role in explaining occupational inequality. Also, income disparity among ethnic groups remains, despite an overall reduction in occupational disparity (see, e.g., Satzewich & Li, 1987; and Rashid, 1985). Indeed, ethnicity may still play an important role in advancement and promotion on the job as well as in the skill requirements of occupations filled by members of different ethnic groups. Relevant to this, Boyd (1985a, p. 401) noted that foreign-born ethnic groups may have difficulties in obtaining recognition for educational credentials in the Canadian labor market and this may limit their life chances. Thus we need to examine the influence of ethnicity on positional advantage within and across occupations.

The implications of the results of this study for educational practice may be stated briefly. Educators are well aware of the tension between the merit and ascriptive criteria in occupational selection and placement and also in the assignment of income. It may be reassuring that the results point to a strong influence of education on occupational status. However, the results also show continued tension between achievement and ascription. The tension between these two principles of allocation is most intense in relation to gender, with women being the more disadvantaged gender group. It also exists, albeit to a much lesser degree, in relation to ethnicity. What this means is that preservice and inservice teacher training programs need to address (or continue to address) gender role and, where necessary, racial biases in the system. As well, school curricula need to be reexamined with a view to correcting these biases. Obviously, educators, school boards, and school teachers and counselors can play a critical role in the move toward gender and ethnic equity.

Similarly, the tension between achievement and ascription is apparent in who gets educated. Educational and hence occupational opportunities are not equally available to all ethnic groups. For example, among the disadvantaged groups are those who are less proficient in the English language. To equalize educational opportunities, school systems need to give higher priority to English as a Second Language (ESL) and other remedial language programs. The emphasis on language proficiency can also be of value to adult immigrants and not only to their children. ESL programs are particularly important because the bulk of immigrants to Canada in the 1990s are projected to be from non-English speaking countries. While such programs may not eliminate ethnic inequalities entirely, they may further weaken the thesis that ethnic affiliation is a correlative factor in the assignment of occupational roles.

Notes

1. Isajiw (1981), on examining language retention of four ethnic groups, has found that retention of ethnic language decreases considerably in the second and subsequent generations.
2. Following Wright's (1977; 1978; 1979) criteria of social class, Li (1988) classifies the "self-employed with paid help" as "employers," and "self-employed without paid help" as "petty bourgeoisie" (p. 68). The rationale for this classification is based on the assumption that both the employers and the petty bourgeoisie own and control money capital and physical capital of production. It appears from this classification that ownership of capital is regarded as a "given" criterion for the employers and the petty bourgeoisie. From this perspective it would be tautological to examine the influence of capital on income.
3. At the initial stage of the analysis socioeconomic status (a composite index of occupation, education, and income) was included but due to multicollinearity between socioeconomic status and education (which is an independent variable included in the analysis), socioeconomic status had to be replaced by occupational status.
4. Of the six ethnic categories, five have been entered as dummy variables with values of 1 in the multiple regression equation. The sixth category, namely, the non-European group, is not included in Table 1 because it has been assigned a value of 0 and treated as a base against which the other categories could be compared. Also, the beta coefficients of French as home language are not included in the table due to multicollinearity between this variable and French ethnic origin on the one hand (.856), and between it and use of English as home language on the other (–.838).
5. Two independent regression analyses were carried out for the British, French, and non-European ethnic groupings, and for those with multiple identity. This was done because of multicollinearity between use of English and use of French as home languages for these four groupings. In the first analysis, use of English as home language was excluded; in the second analysis, use of French as home language was excluded. For the ethnic groupings labeled British, non-European, and multiple identity, use of English as home language was reported because the regression coefficient of this variable was higher than the one associated with use of French as home language. For the French ethnic group use of English as home language was found to be statistically insignificant, and as a result only the analysis that included use of French as home language was reported.

References

Abella, R.S. (1984). *Equality in employment: A royal commission report*. Ottawa: Minister of Supply and Services.

Blau, P.M., & Duncan, O.D. (1967). *The American occupational structure*. New York: Wiley.

Blishen, B.R. (1970). Social class and opportunity in Canada. *Canadian Review of Sociology and Anthropology, 7*, 110-127.

Boulet, J., & Veltman, C. (1981). Socioeconomic achievements of Montreal language groups in 1971. *Canadian Review of Sociology and Anthropology, 18*, 239-248.

Boyd, M. (1985a). Immigration and occupational attainment in Canada. In M. Boyd, J. Goyder, F.E. Jones, H.A. McRoberts, P.C. Pineo, & J. Porter (Eds.), *Ascription and achievement: Studies in mobility and status attainment in Canada* (pp. 393-445). Ottawa: Carleton University Press.

Boyd, M. (1985b). Educational and occupational attainments of native-born Canadian men and women. In M. Boyd, J. Goyder, F.E. Jones, H.A. McRoberts, P.C. Pineo, & J. Porter (Eds.), *Ascription and achievement: Studies in mobility and status attainment in Canada* (pp. 229-296). Ottawa: Carleton University Press.

Boyd, M. (1990). Sex differences in occupational skill: Canada, 1961-1986. *Canadian Review of Sociology and Anthropology, 27*, 285-315.

Boyd, M., Goyder, J., Jones, F.E., McRoberts, H.A., Pineo, P.C., & Porter, J. (1981). Status attainment in Canada: Findings of the Canadian mobility study. *Canadian Review of Sociology and Anthropology, 18*, 657-673.

Cuneo, C., & Curtis, J. (1975). Social ascription in the educational and occupational status attainment of urban Canadians. *Canadian Review of Sociology and Anthropology, 12*, 6-24.

Darroch, A.G. (1980). Another look at ethnicity, stratification and social mobility in Canada. In J.E. Goldstein & R.M. Bienvenue (Eds.), *Ethnicity and ethnic relations in Canada* (pp. 203-230). Toronto: Butterworths.

Duncan, B., & Duncan, O.D. (1968). Minorities and the process of stratification. *American Sociological Review, 33*, 356-364.

Forcese, D. (1980). *The Canadian class structure*. Toronto: McGraw-Hill Ryerson.

Fox, B., & Fox, J. (1986). Women in the labour market 1931-81: Exclusion and competition. *Canadian Review of Sociology and Anthropology, 23*, 1-21.
Goldlust, J. & Richmond, A.H. (1974). A multivariate model of immigrant adaptation. *International Migration Review, 8*, 193-225.
Isajiw, W.W. (1981). *Ethnic identity retention.* Research Paper No. 125, Centre for Urban and Community Studies, University of Toronto.
Kenneth, C.W. (1970). *An introduction to population.* Chicago: Dorsey Press.
Lautard, E.H. & Loree, D.J. (1984). Ethnic stratification in Canada, 1931-1971. *Canadian Journal of Sociology, 9*, 333-344.
Li, P.S. (1978). The stratification of ethnic immigrants: The case of Toronto. *Canadian Review of Sociology and Anthropology, 15*, 30-40.
Li, P.S. (1988). *Ethnic inequality in a class society.* Toronto: Wall and Thompson.
McRoberts, H.A. (1985). Language and mobility: A comparison of three groups: The effects of origin. In M. Boyd, J. Goyder, E.F. Jones, H.A. McRoberts, P.C. Pineo, & J. Porter (Eds.), *Ascription and achievement: Studies in mobility and status attainment in Canada* (pp. 335-356). Ottawa: Carleton University Press.
McRoberts, H.A., Porter, J., Boyd, M., Goyder, J., Jones, F., & Pineo, P. (1976). Differencés dans la mobilité professionnelle des Francophones et des Anglophones. *Sociologie et Sociétés, 8*, 61-79.
Ornstein, M.D. (1981). The occupational mobility of men in Ontario. *Canadian Review of Sociology and Anthropology, 18*, 183-215.
Pedhazur, E.J. (1982). *Multiple regression in behavioral research* (2nd Ed.). New York: Holt, Rinehart & Winston.
Pineo, P.C. (1976). Social mobility in Canada: The current picture. *Sociological Focus, 9*, 109-123.
Pineo, P.C. (1985). Family size and status attainment. In M. Boyd, J. Goyder, F.E. Jones, H.A. McRoberts, P.C. Pineo, & J. Porter (Eds.), *Ascription and achievement: Studies in mobility and status attainment in Canada* (pp. 201-229). Ottawa: Carleton University Press.
Pineo, P., & Porter, J. (1985). Ethnic origin and occupational attainment. In M. Boyd, J. Goyder, F.E. Jones, H.A. McRoberts, P.C. Pineo, and J. Porter (Eds.), *Ascription and achievement: Studies in mobility and status attainment in Canada* (pp. 357-391). Ottawa: Carleton University Press.
Pineo, P.C., Porter, J., & McRoberts, H.A. (1977). The 1971 census and the socioeconomic classification of occupations. *Canadian Review of Sociology and Anthropology, 14*, 91-102.
Porter, J. (1965). *The vertical mosaic.* Toronto: University of Toronto Press.
Rashid, A. (1985). *The Muslim Canadians: A profile.* Ottawa: Minister of Supply and Services.
Reitz, J. (1980). *Survival of ethnic groups.* Toronto: McGraw-Hill Ryerson.
Richmond, A.H. (1967a). *Postwar immigrants in Canada.* Toronto: University of Toronto Press.
Richmond, A.H. (1967b). *Immigrants and ethnic groups in metropolitan Toronto.* Toronto: Institute for Behavioural Research, York University.
Richmond, A.H. (1973). Social mobility of immigrants. In J.E. Curtis & W.G. Scott (Eds.), *Social stratification: Canada* (pp. 174-186). Scarborough: Prentice-Hall of Canada.
Richmond, A.H., & Kalbach, W.E. (1980). *Factors in the adjustment of immigrants and their descendants.* (Cat. No. 99-761). Ottawa: Minister of Supply and Services.
Sandefur, G.D., & Scott, W.J. (1981). A dynamic analysis of migration: An assessment of the effects of age, family and career variables. *Demography, 18*, 355-368.
Satzewich, V., & Li, P.S. (1987). Immigrant labour in Canada: The cost and benefit of ethnic origin in the job market. *Canadian Journal of Sociology, 12*, 229-241.
Sharma, R.D. (1980). *Migration and fertility in a western Canadian metropolis.* Unpublished doctoral dissertation, University of Alberta.
Tepperman, L. (1975). *Social mobility in Canada.* Toronto: McGraw-Hill Ryerson.
Trovato, F., & Grindstaff, C.F. (1986). Economic status: A census of immigrant women at age thirty in Canada. *Canadian Review of Sociology and Anthropology, 23*, 569-587.
Wright, E.O. (1977). Marxist class categories and income inequality. *American Sociological Review, 42*, 32-55.
Wright, E.O. (1978). Race, class and income inequality. *American Journal of Sociology, 83*, 1368-1397.
Wright, E.O. (1979). *Class structure and income determination.* New York: Academic Press.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 3, September 1992, 219-234*

Estelle S. Gellman
Hofstra University

and

Mina Berkowitz
Long Island University

Factors Perceived as Important in Teacher Evaluation

The study sought to determine the factors that different segments of the school community consider important in evaluating a teacher. Elementary and secondary school teachers, novice and experienced teachers, student teachers, teacher educators, and parents were asked to rate the importance of 39 teacher characteristics. A factor analysis of their responses yielded three factors: (a) a responsiveness/attitude toward students factor, (b) a discipline/classroom management factor, and (c) a subject-matter expertise factor. Multivariate analyses of variance indicated that the responsiveness/attitude toward students factor was consistently rated as most important. The experienced teachers, the teacher educators, and the elementary school teachers rated the subject-matter expertise factor as significantly more important than the discipline/classroom management factor, while the reverse was true for the student teachers. No significant differences were found in the importance assigned to these two factors by the other groups. The findings indicate a common basis for teacher evaluation.

L'étude a tenté de déterminer les facteurs que différents segments de la communauté scolaire considèrent comme importants dans l'évaluation d'un maître. On a demandé à des maîtres des écoles élémentaires et secondaires, des maîtres novices et des maîtres expérimentés, des élèves-maîtres, des formateurs de maîtres, et des parents d'évaluer l'importance de 39 caractéristiques des maîtres. Une analyse factorielle de leurs réponses a révélé trois facteurs: (a) un facteur de réponse/attitude positive envers les élèves, (b) un facteur de discipline/gestion de classe et (c) un facteur d'expertise dans la matière d'enseignement. Des analyses multivariées de la variance ont indiqué que le facteur de réponse/attitude positive envers les élèves était toujours considéré comme le plus important. Les maîtres expérimentés, les formateurs de maîtres et les maîtres des écoles élémentaires ont évalué le facteur d'expertise dans la matière d'enseignement comme plus important, et cela d'une manière significative, que le facteur de discipline/gestion de classe, ce qui était le contraire pour les élèves-maîtres. Il n'y avait pas de différences significatives dans l'importance accordée à ces deux facteurs par les autres groupes. Les résultats indiquent une base commune pour l'évaluation des maîtres.

Estelle Gellman is a professor in the Department of Counseling, Research, Special Education and Rehabilitation where she teaches graduate courses in educational research, measurement, and statistics.
Mina Berkowitz is a professor of education at the Brooklyn Campus where she teaches courses in measurement and educational psychology in the undergraduate program in elementary education and the graduate programs in school psychology, special education, and elementary education.

Although there has been much interest in the importance of attracting good teachers into the classroom, there is not necessarily agreement as to what constitutes a "good" teacher (Epstein, 1985; Hamacheck, 1969; Issler, 1983; Tolor, 1973). Notwithstanding the large body of research investigating the relationship between various teacher behaviors and student achievement (Brophy & Good, 1986), the question remains as to whether the behaviors that result in improved student achievement are those that are perceived as characteristic of a good teacher. Furthermore, there may be different perceptions of good teaching among different segments of the school community. As indicated by Darling-Hammond, Wise, and Pease (1983), different evaluation constituencies—teachers, principals, parents, and public officials—have different stakes in the evaluation process. There is some indication that these different constituencies do in fact have different outlooks on the teaching profession and that these differences are reflected in differing perceptions of what constitutes a good teacher. It has been found, for example, that parent ratings of teacher effectiveness differ from those of faculty members and administrators (Epstein, 1985; Tolor, 1973). It may well be that even professional educators are not in agreement about what constitutes a good teacher. In a study investigating the importance that principals and secondary school biology teachers attach to different criteria of teacher effectiveness, Searles and Ng (1982) found that characteristics dealing with subject-matter preparation and enthusiasm for the subject were among the most highly rated criteria by both groups. In Epstein's (1985) study of elementary school teachers, on the other hand, it was found that neither principals' nor parents' ratings of teachers were correlated with advanced study in the subjects taught. These differences suggest that there may be systematic variation in the way that different segments of the educational community conceptualize a good teacher. If there are different perceptions of good teaching, questions may arise as to the criteria to be used in assessing teacher performance.

In reviewing the literature on teacher evaluation, Strike and Millman (1983) suggest that, in order for assessment to be successful, teachers and evaluators must have common views concerning education, evaluation, and the role of the teacher. Kauchak, Peterson, and Driscoll (1985) further suggest that professionalism demands that teacher perspectives be considered in any evaluation system. Selmes (1986) goes so far as to suggest that teacher evaluation should be based on a new set of assumptions, including the assumption that evaluation criteria should be jointly negotiated and agreed upon by teachers and evaluators.

If successful evaluation procedures require agreement on the criteria of evaluation, then it is important to know the extent to which different segments of the school community agree on the characteristics to be assessed. Different perceptions of a good teacher may be held by elementary and secondary school teachers, teachers differing in their degree of experience, student teachers, and teacher educators. Parents and teachers may also differ in their perceptions of good teaching. These different groups may differ not only in those characteristics that are identified as being important, but also in the degree of importance ascribed to particular characteristics.

In order to explore these issues, the aim of the present study was to first identify the factors perceived as important in evaluating a teacher. Although the literature on teacher effectiveness identifies a variety of discrete behaviors related to student outcomes, the purpose of the present study was to determine whether in examining perceptions of good teaching these characteristics could be explained by fewer more general factors. That such is often the case is indicated by factor analytic studies of teacher evaluation instruments. Swartz, White, Stuck, and Patterson (1990), for example, in examining the factorial structure of 27 teaching practices on the North Carolina Teaching Appraisal Instrument (TPAI), found that the items on the scale were most parsimoniously described by two factors, one relating to the teacher's presentation of the lesson and the other relating to the teacher's management of the class. Similarly, Teddlie, Virgilio, and Oescher (1990) found that the 35 items on the Virgilio Teacher Behavior Inventory (VTBI) could be explained by five factors: one classroom management, two instruction, and two classroom climate dimensions.

The second part of the present study then examined the importance ascribed to each of the emerging factors. The overall importance rating of each factor was examined as well as the relative importance ascribed to each of the factors by different segments of the school community: parents, teachers (broken into subgroups of novice, experienced, elementary, and secondary teachers), student teachers, and teacher educators.

Method

In order to identify those factors that different groups perceive as important in evaluating a teacher, a questionnaire was developed that asked respondents to indicate the importance that they would ascribe to each of 39 teacher characteristics.

Construction of the Instrument

A preliminary list of teacher characteristics was derived from a review of the literature on teacher evaluation. Additional characteristics were then obtained by asking teachers, parents attending PTA meetings, and high school students attending classes at two school districts in Nassau County to list the five characteristics that they would consider most important in evaluating a teacher. The items were originally grouped into four categories: (a) classroom management/climate, (b) personal qualities, (c) subject matter competence/communication skills and (d) interpersonal relationships. Feedback from a small group of educators indicated, however, that these categories were insufficient; several of the items did not appear to be appropriately placed in any of these categories. The experimenters then reclassified the items and assigned category names to similarly grouped items. It was felt that the items were best described by the following seven category names: (a) discipline, (b) classroom management, (c) subject matter competence, (d) attitude toward students, (e) demands on students, (f) pedagogical skills, and (g) personal characteristics.

A preliminary scale was then developed which included between seven and 15 items from each category. Where a category had fewer than seven items, additional items were written by the investigators. These items were

then distributed to 158 teachers and graduate education students with instructions to indicate the category represented by each item.

A seven-point Likert-type scale of 39 items was then developed that included those items (a) on which there was over 48% agreement on the category assignment, and (b) that were not assigned to any other category more than 32% of the time. The mean percent agreement for the items in each category ranged from 54.43% to 92.7%, with a mean for the entire scale of 75%. Three forms of the final scale were constructed, each form containing a different random order of the items. Respondents were instructed to indicate the degree of importance that they would assign to each of the 39 characteristics in evaluating a teacher by circling the appropriate scale value. Values ranged from 1 through 7, with a value of 1 indicating that the characteristic was considered a major factor in evaluating a teacher and a value of 7 indicating that the characteristic was considered to be totally irrelevant.

Sample and Procedure

The scale was administered to 488 parents and educators in the New York metropolitan area, with random assignment of the different forms. Subjects were sought at PTA meetings, professional meetings, graduate education classes, and university departments of education. Respondents were assured of anonymity and were requested to complete a personal data form as well as to rate the importance of each of the items on the scale of teacher characteristics.

As noted in Table 1, the sample consisted of 102 males and 374 females, with 12 respondents not completing this item. Included in the sample were 30 teacher educators, 53 student teachers, and 64 parents who were not teachers. The sample also included 120 elementary school teachers, defined as those

Table 1
Number of Respondents in each Category Broken Down by Sex and Experience

		Years of Experience														
		1-5			6-10			11 or more			Missing			Total		
Category	n	M	F	NR	M	F	NR	M	F	NR	M	F	NR	M	F	NR
Elementary Teachers	120	4	45	0	2	20	0	5	19	0	3	20	2	14	104	2
Secondary Teachers	100	11	20	0	9	10	0	20	14	2	3	11	0	43	55	2
Teacher Educators	30	0	1	0	1	1	0	9	10	0	2	5	1	12	17	1
Student Teachers	53	—	—	—	—	—	—	—	—	—	—	—	—	5	45	3
Parents	64	—	—	—	—	—	—	—	—	—	—	—	—	9	54	1
Total Sample	488	—	—	—	—	—	—	—	—	—	—	—	—	102	374	12

Note. NR represents individuals who did not indicate their sex. The subgroup totals do not equal the total sample because not all individuals could be assigned to a subgroup. Dashes indicate the absence of respondents in the category.

who taught in grades N-6 and had no junior high school, high school or college teaching experience, and 100 secondary school teachers, defined as those who taught in grades 7-12 and had no experience at the elementary or nursery school level. Teachers who had experience at both the elementary and secondary levels were not included in either of these subgroups. In the sample of teachers, 121 were new teachers, defined as having had five or fewer years of experience, and 108 were experienced teachers, defined as having had 11 or more years of experience.

Data Analysis

Part 1. Part 1 of the study was designed to identify the factors that are perceived as important in evaluating a teacher.

An initial iterative principal axis factor analysis was performed on the importance ratings of those respondents who completed all of the items on the scale. The purpose of this analysis was to ascertain how many factors would best describe the data. The criteria entering into this decision included (a) the initial eigenvalues (i.e., not subjected to iteration), (b) a scree examination, and (c) the percent of variance explained by the factor. Based on these criteria, the three- and four-factor solutions were further examined for interpretability and were subjected to a principal axis factor analysis with a Varimax rotation. This analysis guided the final judgment as to the number of factors that best explained the data and was used for the development of factor-based scales (Kim & Mueller, 1978, p. 70). SPSS, Release 4.0 for the IBM VM/CMS, was used for all data analyses.

Part 2. Part 2 of the study examined the relative importance assigned to the factors identified in Part 1.

In order to compare the importance ascribed to the factors by the different groups, a factor-based scale was developed for each factor. For an item to be included on a scale, it had to load .4 or higher on the rotated factor in the three-factor solution and have a loading less than .35 on the other two factors.

Based on the scores on the factor-based scales, a repeated measures multivariate analysis of variance was performed to compare the importance ascribed to each of the factors by the total sample. Then four profile analyses were performed to compare the ratings assigned by different subgroups to the different factors. The groups compared were (a) elementary and secondary school teachers, (b) new and experienced teachers, (c) parents and experienced teachers, and (d) student teachers, teacher educators, and experienced teachers.

Results

Part 1

As reported in Table 2, the initial factor analysis yielded four factors with eigenvalues greater than 1, with the fourth factor barely meeting this criterion. A scree examination yields an "elbow" that falls to the left of the fourth factor, with the eigenvalues clearly leveling off by the fourth factor. As may be also noted in Table 2, the first three factors explained 79.21% of the variance, with the fourth factor explaining another 6.92%.

Based on the above, both the three- and the four-factor solutions were further examined for interpretability. Inspecting the rotated factor matrices, as

Table 2
Initial Eigenvalues and Percent of Variance Explained by Each Factor in the Iterative Factor Analysis

Factor	*Eigenvalue*	*Percent of Variance Explained*
1	7.762	49.55
2	2.730	17.43
3	1.916	12.23
4	1.083	6.91
5	.976	6.23
6	.702	4.48
7	.674	4.30
8	.491	3.13
9	.441	2.82
10	.356	2.27
11	.283	1.81
12	.254	1.62
13	.242	1.54
14	.178	1.14
15	.143	.91
16	.121	.77
17	.100	.64
18	.069	.44
19	.053	.34
20	.022	.14
21	.006	.04
22	−.003	<.01
23	−.023	<.01
24	−.054	<.01
25	−.081	<.01
26	−.100	<.01
27	−.102	<.01
28	−.133	<.01
29	−.141	<.01
30	−.154	<.01
31	−.179	<.01
32	−.196	<.01
33	−.204	<.01
34	−.230	<.01
35	−.240	<.01
36	−.246	<.01
37	−.263	<.01
38	−.275	<.01
39	−.314	<.01

reported in Table 3, it may be seen that only five items had factor loadings of .4 or higher on the fourth factor and that none of them had loadings greater than .61. Furthermore, one of these items (item 6) had a higher loading on Factor 2, and two of the items had loadings in the area of .30 on either Factor 1 or Factor 2.

Looking further at the item correlation matrix, it was found that the intercorrelations among the items with high loadings on the fourth factor ranged from .17 to .47, with a mean of .31. Many of the correlations between the items

Table 3
Rotated Factor Matrices for the Three and Four-Factor Solutions

	Factor Loading						
	Three Factor Solution			Four Factor Solution			
Item	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 4
1	.37787	.12822	.05616	.29351	.03029	.04155	.43555
2	.03216	.20330	.37192	.03467	.20968	.37103	.02006
3	.54809	−.06066	.01172	.54761	.05486	.00996	.06112
4	.25341	.11860	.54135	.22137	.08851	.53754	.18337
5	.20224	.60761	.04420	.15928	.56919	.03586	.24887
6	.26692	.51790	.15474	.17896	.43854	.14062	.43644
7	.21181	.06772	.60636	.23047	.09260	.61053	−.02949
8	.60375	−.04192	.03989	.59663	−.04228	.03721	.09546
9	.06219	.22293	.15342	.05184	.21475	.15146	.07127
10	.00763	.11663	.51102	.00199	.11237	.50812	.04980
11	.07561	.62401	.11678	.07801	.63982	.11304	.04980
12	.54125	.04628	.01246	.52594	.03745	.00824	.12789
13	.56643	.07550	.13017	.57631	.09283	.12845	.04171
14	.06421	.49383	.12093	.03084	.46278	.11532	.18617
15	−.02195	.14716	.01376	−.00627	.16575	.01589	−.05474
16	.51721	.20751	.14141	.48230	.17847	.13486	.22097
17	−.01390	.53533	.04573	−.03197	.51939	.04213	.11824
18	.55789	.03356	.05478	.54619	.02775	.05098	.11730
19	.14747	.69800	.05105	.13257	.69209	.04484	.13619
20	.37410	.18125	.13908	.27016	.06111	.12605	.53259
21	.61560	.17338	.06788	.57830	.14249	.06045	.23772
22	.30035	.46504	.19576	.30749	.48523	.19335	.04993
23	.42487	.21420	.21462	.46066	.25996	.21766	−.05908
24	.41924	.09317	.09300	.39677	.07501	.08882	.14896
25	.40021	.20341	.12043	.42256	.23480	.12019	−.01676
26	.34906	.31430	.20562	.22670	.18816	.19533	.60529
27	.62666	.11790	.08445	.62495	.12526	.08105	.09383
28	.49068	.40640	.08497	.48884	.41626	.08029	.10289
29	.68557	.09135	−.04066	.65535	.06819	−.04672	.20642
30	.05984	.55726	.19792	.06296	.57167	.19571	.04214
31	.27614	.39637	.11417	.18170	.30715	.09944	.45845
32	.55723	.14273	−.08793	.52726	.11911	−.09358	.19315
33	.52809	−.03208	.10834	.57089	.01272	.11151	−.08880
34	.18041	.07226	.68187	.18065	.07764	.67882	.04538
35	.11872	.62133	.01942	.12142	.63800	.01471	.05194
36	.50744	.13048	.16616	.49601	.12684	.16278	.11817
37	.02278	.08620	.74299	.02070	.08724	.73993	.03945
38	−.00695	.10324	.45496	−.05296	.05503	.45579	.20096
39	.17137	.20968	.18934	.20967	.25930	.19436	−.11265

with high loadings on Factor 4 and the items with high loadings on Factor 1 or Factor 2, however, fell in the same range. Additionally, as may be noted in Table 3, every item on Factor 4 had a loading greater than .30 on either Factor 1 or Factor 2 in the three-factor solution, with item 6 having a higher loading on Factor 2 in the three-factor solution than it did on Factor 4.

Considering the substance of the items with high loadings on Factor 4, they might be broadly interpreted as representative of a personal qualities factor,

including such items as (a) Has a good personality, (b) Has a pleasant voice, and (c) Has a sense of humor. These items, however, do not seem to meaningfully fit with item 31, Displays good use of bulletin board space, which has the third highest loading on the factor. In the three-factor solution, the factor loading of item 31 increases on Factor 2, which appears to be related to classroom management, and the other items increase their loadings on Factor 1, which appears to reflect the teacher-student relationship. Taking these considerations into account, as well as the question of the reliability of a factor-based scale including those few high-loading items on Factor 4, only three factors were used for subsequent analyses. The items included on the three factor scales are presented in Table 4. Nine of the items on the original scale were not included on any of the factor-based scales.

Those items with high factor loadings on each of the factors were examined in order to define the three factors. Factor 1 was defined as a responsiveness/attitude toward students factor, Factor 2 as a discipline/classroom management factor, and Factor 3 as a subject-matter expertise factor.

Table 4
Items Included on each Factor Scale

Factor 1	*Factor 2*	*Factor 3*
Respects the students (#3)	Attendance is taken regularly (#5)	Reads extensively in the subject area (#4)
Listens to students (#8)	Is neatly dressed (#6)	Knows subject in depth (#7)
Is patient (#12)	Is a strict disciplinarian (#11)	Is an acknowledged expert in the subject taught (#10)
Admits mistakes honestly (#13)	Does not allow talking during instruction (#14)	Has taken advanced courses in subject area (#34)
Asks questions in class (#16)	Students are punished for being late (#17)	
Is enthusuastic (#18)	Maintains a strict set of rules to be obeyed (#19)	Has an advanced degree in the subject area (#37)
Uses a variety of teaching skills (#21)	Keeps accurate records (#22)	Has professional publications in subject area (#38)
Gives fair exams (#23)	Regularly gives quizzes and exams (#30)	
Respects students' privacy (#24)	Misbehavior is swiftly punished (#35)	
Sets demands on level of students (#25)		
Provides feedback (#27)		
Makes students feel important (#29)		
Knows each student's level of competence (#32)		
Stimulates intellectual curiosity (#33)		
Repeats material as needed (#36)		

Part 2.

Total group analysis. When looking at the mean importance ratings assigned to the three factors by the total group, as reported in Table 5, it was found that the responsiveness/attitude toward students factor was rated as more important than the subject-matter expertise factor, and that the subject-matter expertise factor was rated as more important than the discipline/classroom management factor. The overall multivariate analysis of variance was significant; Wilks' λ=.189, $F(2,482)$=1,033.55, $p<.05$. It may also be noted, as reported in Table 6, that the difference in the importance ascribed to the responsiveness/attitude toward students factor and the subject-matter expertise factor was also significant, as was the difference between the importance ratings of the subject-matter expertise factor and the discipline/classroom management factor.

Subgroup analyses. As also noted in Table 5, every subgroup rated the responsiveness/attitude toward students factor as most important and, as may be further noted in Table 6, the responsiveness/attitude toward students factor was rated by every group as significantly more important than the next highest rated factor.

The subject-matter expertise factor was rated as more important than the discipline/classroom management factor by all subgroups except the parents and the student teachers. As may be noted in Table 6, the difference in the importance ratings of these two factors was significant for all subgroups except for the: (a) new teachers, (b) secondary school teachers, and (c) parents.

Examining the pattern of sub-group response across the factors, the profile of the elementary school teachers was significantly different from that of the

Table 5
Mean Importance Ratings Assigned to the Three Factors by the Different Groups

		Factor					
		Responsiveness to Students		*Discipline/ Management*		*Subject-Matter Expertise*	
Group	*n*	*Mean*	*Standard Deviation*	*Mean*	*Standard Deviation*	*Mean*	*Standard Deviation*
Parents	64	1.75	.53	3.03	.91	3.22	.94
Experienced Teachers (All)	108	1.85	.62	3.69	1.14	3.15	.97
Experienced Teachers (Excluding teacher educators)	89	1.94	.63	3.77	1.17	3.19	.98
New Teachers	121	1.71	.45	3.38	.85	3.22	.95
Elementary Teachers	118	1.71	.46	3.61	1.00	3.31	1.00
Secondary Teachers	100	1.79	.51	3.30	1.05	3.27	.92
Teacher Educators	29	1.51	.36	3.52	.98	3.07	.91
Student Teachers	52	1.71	.53	3.23	.82	3.70	1.02
Total Sample	484	1.75	.54	3.42	1.00	3.26	.97

Note. Where n is less than the number of such individuals in the sample, it is because respondents who did not respond to every item on the scale were not included in this analysis.

Table 6
Significance of the Difference in the Importance Ratings Assigned to the Three Factors by the Groups Compared in the Profile Analyses

			Comparison					
			Factor 1 vs Factor 2		*Factor 1 vs Factor 3*		*Factor 2 vs Factor 3*	
Group	n	df	F	*p*	F	*p*	F	*p*
Total Sample	484	1,483	1,478.53	.000	1,128.71	.000	8.37	.004
Teacher Educators	29	1,167	121.27	.000	77.91	.000	4.95	.027
Student Teachers	52	1,167	126.07	.000	227.54	.000	9.75	.002
Experienced Teachers (Excluding Teacher Educators)	89	1,167	309.36	.000	152.79	.000	25.93	.000
Experienced Teachers (Including Teacher Educators)	108	1,227	392.29	.000	187.02	.000	22.24	.000
New Teachers	121	1,227	365.26	.000	287.22	.000	2.15	.144
Elementary Teachers	118	1,216	440.71	.000	330.02	.000	7.28	.008
Secondary Teachers	100	1,216	236.31	.000	239.55	.000	.08	.779
Parents	64	1,170	101.68	.000	155.86	.000	2.07	.152
Experienced Teachers (Including Teacher Educators)	108	1,170	356.55	.000	205.08	.000	28.21	.000

Note. Where n is less than the number of such individuals in the sample, it is because respondents who did not respond to every item on the scale were not included in this analysis.

secondary school teachers (λ=.962, $F(2,215) = 4.27$, $p<.05$). Although both groups rated subject-matter expertise as more important than discipline/classroom management, the difference in the ratings of these two factors was significant for the elementary school teachers but not the secondary school teachers. It may also be noted that the profile of the parents was significantly different from that of the experienced teachers (λ=.892, $F(2,169)=10.21$, $p<.05$). While the experienced teachers rated subject-matter expertise as significantly more important than discipline/classroom management, this pattern was not duplicated for the parents; although the parents reversed the relative importance ascribed to discipline/classroom management and subject-matter expertise, the difference between these two factors was not significant for the parents.

A significant difference was also found in the profile analysis comparing the student teachers, teacher educators and experienced teachers (λ=.820, $F(4,334)= 8.67$, $p<.05$). The analysis did not indicate any significant difference among the groups in the degree to which Factor 1 was seen as more important than Factor 2 ($F(2,167)=2.59$, $p.>05$). Differences were found, however, in the differentiation between Factors 2 and 3 ($F(1,167)=16.18$, $p<.05$). Contrasts indicated that although the experienced teachers and the teacher educators did not

significantly differ in the extent to which they differentiated among Factors 2 and 3 ($F(1,167)=.35$, $p>.05$), there was a significant difference between these two groups and the student teachers ($F(1,167)=26.98$, $p<.05$). As may be noted in Table 5, this difference reflects the fact that, while the student teachers rated the discipline/classroom management factor as significantly more important than the subject-matter expertise factor, the reverse was true for the experienced teachers and teacher educators.

No significant profile differences were found when comparing the importance ascribed to the three factors by the new and experienced teachers.

Discussion

Although the present study was based on responses to a questionnaire designed to include items relating to seven different aspects of teaching cited in the literature, the factor analysis did not yield seven distinct factors. Only three clear factors were found to underlie the perception of what constitutes a good teacher: (a) a responsiveness/attitude toward students factor, (b) a discipline/classroom management factor, and (c) a subject-matter expertise factor.

The responsiveness/attitude toward students factor emerged first and was most diffuse. This factor seems to reflect aspects of both personal and pedagogical responsiveness, as well as general personal and attitudinal characteristics. Most of the high-loading items on this factor had been assigned to four of the seven categories on which the questionnaire was based: (a) attitude toward students, (b) demands on students, (c) pedagogical skills, and (d) personal characteristics. It is quite clear, however, that these items did not reflect four distinct factors.

The discipline/classroom management factor was the second factor to emerge. In constructing the survey instrument, we assigned most of the items with high loadings on this factor to what appeared to be two distinct categories, discipline and classroom management. The results of the factor analysis indicated, however, that items from both categories were highly intercorrelated and had high loadings on a single factor. This factor was much more sharply defined than the first factor, with only one high-loading item, "Is neatly dressed," not clearly reflecting a classroom management and/or discipline function. One might, however, interpret this item to be reflective of self-discipline on the part of the teacher.

The third factor was also very sharply defined, with every high-loading item clearly reflecting advanced training or expertise in the subject-matter being taught. Every item with a high loading on this factor had been originally assigned to the category of subject-matter expertise, with only one questionnaire item originally assigned to this category not meeting the requirements for inclusion on the factor-based scale for the third factor. This item, "Is able to answer most questions without a book," had a factor loading of .37 on the factor.

The factors emerging in the present study are not surprising in view of the Swartz et al. (1990) and Teddlie et al. (1990) studies of teacher appraisal instruments. Although the factor structure is not identical to that found in the other studies, there is considerable similarity even though these other studies investigated the structure of the practices assessed on actual appraisal instru-

ments whereas the present study focused on the importance ascribed to various teacher behaviors.

Classroom management, for example, emerged as a single factor in the present study as well as in both the Teddlie et al. (1990) and the Swartz et al. (1990) studies. Of particular interest is the fact that, although the TPAI (Swartz et al., 1990) yields separate scores for management of instructional time and management of student behavior, the investigators concluded on the basis of a factor analysis of teacher ratings that a two-factor solution in which aspects of both types of behavior loaded on a single factor was most parsimonious in describing the data. The instrument constructed for the present study was also designed with separate categories for discipline and classroom management, and the results of the factor analysis similarly indicated that these practices formed a single factor.

The second factor emerging from the analysis of the TPAI was defined as an instructional presentation factor on which the high loading practices parallel many of the high-loading items found on the responsiveness/attitude toward students factor in the present study. The other factors emerging from the analysis of the VTBI (Teddlie et al., 1990) included two instruction factors and two classroom climate factors. Although the present study did not find distinct factors that were definable in this way, many of the items with high loadings on Factor 1 in the present study do reflect classroom climate or instruction. The question is whether these practices can be described by a single underlying factor as found in the present study and the Swartz et al. (1990) study or whether more factors are needed to meaningfully describe the data.

It is also of interest to note that neither the Swartz et al. (1990) nor the Teddlie et al. (1990) studies yielded anything reflecting the subject-matter expertise factor emerging in the present study. This finding is understandable because both the TPAI and the VTBI are essentially observation instruments that focus on observable teacher behaviors. The factors emerging from factor analytic studies of any assessment instrument are necessarily restricted by the nature of the items being assessed, and a subject-matter expertise factor would not emerge in a study where this factor is clearly not assessed. The present study included items relating to this variable, and these items clearly formed a factor distinct from the discipline/classroom management and responsiveness attitude toward students factor.

Interestingly, these latter two factors are also congruent with the research on teacher effectiveness. The research of Soar and Soar (cited in Brophy & Good, 1986), for example, indicates that emotional climate and teacher management are independent factors affecting achievement. Their definition of emotional climate includes many of the characteristics included on the first factor emerging in the present study, whereas their definition of teacher management subsumes both the classroom management and discipline aspects of Factor 2. In the Brophy and Good review (1986), in fact, classroom management was repeatedly cited as a factor related to student achievement and, although the findings are more mixed, several of the studies reviewed also found student achievement to be related to teacher behaviors that had high

loadings on the responsiveness/attitude toward students factor emerging in the present study.

The difference between the factors that were cited in the Brophy and Good review (1986) as related to student achievement and those that the present study found to be considered important in teacher evaluation was that (a) the subject-matter competence of the teacher was not cited as a factor in the studies on student achievement, and (b) many of the specific pedagogical techniques found to be related to student achievement were not included on any of the factors in the present study. Although the questionnaire used in the present study did include general items relating to such areas as feedback, questioning, demands on students, and setting levels of expectation, it did not include items relating to specific classroom techniques such as grouping, rephrasing, wait time, time on task, pacing, providing an overview or summary, et al.

Those items that did relate to more general pedagogical techniques tended to load on the responsiveness/attitude toward students factor. It would appear that such teacher behaviors as questioning, providing feedback, and adjusting demands to the level of the students are perceived within the general factor of responsiveness to student learning needs. Because the present study did not include the more specific behaviors, it is not clear whether they also would be seen in this way. Further investigation is needed to clarify this question, but the results of the present study suggest that many of the pedagogical practices related to student achievement may have a common thread of reflecting a responsiveness to student needs. It seems that this factor reflects more than positive affect on the part of the teacher and also includes a recognition of the cognitive needs of the student.

The third factor emerging in the present study, subject-matter expertise, has been only tangentially mentioned in the research on teacher effectiveness. One explanation may be that the research in this area parallels assessment instruments such as the TPAI and the VTBI in focusing on observable teacher behaviors. Issues such as previous schooling or involvement with out-of-class professional activities are not directly observable, and it may be assumed that these factors are a prerequisite to such variables as clarity of presentation or fluency. The present study does indicate, however, that subject-matter expertise, whether assumed to be prerequisite to other teacher behaviors or not, is a separate and distinct factor to be considered in evaluating a teacher.

In examining the importance ascribed to the three factors emerging in the present study, it is of particular interest to note the high degree of consistency among the different segments of the school community in rating the responsiveness/attitude toward students factor as most important. Not only was the importance ascribed to the discipline/classroom management and subject-matter expertise factors consistently lower than that ascribed to the responsiveness/attitude toward students factor, but in every instance these two factors were also seen as closer in importance to each other than either was to the responsiveness/attitude toward students factor.

Although the analysis of the responses of the total group indicated that significantly more importance was ascribed to the subject-matter expertise factor than to the discipline/classroom management factor, this pattern was

not consistent among the subgroups. Whereas the teacher educators, experienced teachers, and elementary school teachers rated the behaviors on the subject-matter expertise factor as more important than those on the discipline/classroom management factor, the reverse was true for the student teachers. The parents, the new teachers, and the secondary school teachers did not rate these two factors as differing significantly in importance.

The consistency of this pattern is borne out in the profile analyses that were performed to compare the pattern of response among student teachers, teacher educators, and experienced teachers and between new and experienced teachers, elementary and secondary teachers, and parents and experienced teachers. In those instances where profile differences were found, the difference could be attributed to the fact that the groups compared differed in the relative importance ascribed to Factors 2 and 3.

Although the general pattern was to rate the subject-matter expertise factor as more important than the discipline/classroom management factor, it is interesting to note that the two groups in which this pattern was not found were the parents and the student teachers, the two groups without any or with minimal teaching experience. Also, there was no significant difference in the importance that the new teachers ascribed to these two factors. One may speculate that experienced teachers who know that they are capable of handling a class of students may be less concerned with issues of discipline than with subject-matter competence where, in order to remain current in the field, they must continuously update their knowledge and skill. Student teachers, on the other hand, are in the process of learning the latest developments in the field, but have not had the experience of testing their own competence in disciplining and managing a class. They may therefore perceive this task as more challenging and see success in this areas as more important.

It is also interesting to note that, while the elementary school teachers rated the subject-matter expertise factor as more important than the discipline/classroom management factor, there was no difference in the importance ascribed to these two factors by the secondary teachers. Given the different teaching environments of elementary and secondary school teachers, one might expect secondary school teachers to focus more on subject-matter expertise. The fact that this factor was not rated as more important may suggest that elementary school teachers assume that most of their colleagues have the minimum management abilities required and have more concern about teachers remaining current in the subject areas being taught than in their classroom management abilities. Secondary school teachers, on the other hand, are dealing with an age group that can be extremely disruptive in the absence of appropriate management and discipline practices and may consequently put as much emphasis on this factor as on subject-matter expertise.

These findings clearly indicate that the way teachers respond to students is perceived as being of primary importance among all segments of the school community, and that the attitude the teacher displays toward students is perceived as more important than either the depth of knowledge in the subject being taught or classroom management skills. It should be noted, however, that all the factors were considered important. The mean on each of the factor scales fell on the positive side of the scale, where a rating of 1 indicated that the

behavior was considered a major factor in teacher evaluation. While the responsiveness/attitude toward students factor consistently fell between the most positive values of 1 and 2 on a 7-point scale, however, the mean of the other two factor scales fell between 3 and 4.

The importance that is ascribed to this factor clearly indicates that teacher evaluation must focus on ways to assess those complex and relatively intangible behaviors that reflect the quality of the interaction between teacher and student and appear to be perceived as the foundation of good teaching. Although this interaction may be difficult to assess and may not be necessarily related to the conventional indices of student achievement, it must be included in any teacher assessment process designed to reflect those aspects of teaching that teachers themselves consider important.

If, as suggested by Strike and Millman (1983), there must be agreement between teacher and evaluator for assessment to be successful, the strong agreement that was found among the different segments of the school community indicates that successful evaluation is possible. According to Darling-Hammond, Wise and Pease (1983), because approaches to teacher evaluation vary according to those teacher attributes believed to be important for effective teaching, it is clear that any approach to evaluation must address the issue of how to assess those less tangible behaviors that define the nature of the teacher's attitude and relationship to students.

Acknowledgments

We gratefully acknowledge the assistance of Liora Schmelkin in analyzing and interpreting the data and of Marci Mann in searching the literature, in recording and entering the data on the computer, and in executing the computer analyses.

Portions of this research were presented at meetings of the Northeastern Educational Research Association and the Eastern Educational Research Association.

References

Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Third handbook of research on teaching* (pp. 328-375). New York: Macmillan.

Darling-Hammond, L., Wise, A.E., & Pease, S.R. (1983). Teacher evaluation in the organizational context. *Review of Educational Research, 53*(3), 285-329.

Epstein, J.L. (1985). A question of merit: Principals' and parents' evaluations of teachers. *Educational Researcher, 14(7), 3-10.*

Hamachek, D. (1969). Characteristics of good teachers and implications for teacher education. *Phi Delta Kappan, 50,* 341-344.

Issler, K. (1983). A conception of excellence in teaching. Education, 103, 338-343.

Kauchak, D., Peterson, K., & Driscoll, A. (1985). An interview study of teachers' attitudes toward teacher evaluation practices. *Journal of Research and Development in Education, 19,* 32-37.

Kim, J., & Mueller, C.W. (1978). *Factor analysis: statistical methods and practical issues* (Sage University paper series on Quantitative Applications in the Social Sciences, 07-014). Beverly Hills: Sage.

Searles, W.E., & Ng, R.W.M. (1982). A comparison of teacher and principal perception of an outstanding biology teacher, *Journal of Research in Science Teaching, 19,* 487-495.

Selmes, A. (1986). Teacher evaluation in the classroom. *Educational Management and Administration, 14,* 191-196.

Strike, K.A., & Millman, J. (1983). Nontechnical questions about teacher evaluation systems in elementary and secondary schools: A research agenda. *Educational Evaluation and Policy Analysis, 5,* 389-397.

Swartz, C.W., White, K.P., Stuck, G.B., & Patterson, T. (1990). The factorial structure of the North Carolina Teaching Performance Appraisal Instrument. *Educational and Psychological Measurement, 50,* 175-182.

Teddlie, C., Virgilio, I., & Oescher, J. (1990). Development and validation of the Virgilio Teacher Behavior Instrument, *Educational and Psychological Measurement, 50*, 421-430.

Tolor, A. (1973). Evaluation of perceived teacher effectiveness. *Journal of Educational Psychology, 64*, 98-104.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 3, September 1992, 235-250*

Dushyanta Persaud
Winnipeg School Division No. 1

and

Paul R. Madak
University of Manitoba

Graduates and Dropouts: Comparing Perceptions of Self, Family, and School Supports

A survey approach was used to compare 92 students about to graduate from three Winnipeg high schools with 38 dropouts. Graduates were obtained from an alternative school, a vocational school and a school geared towards academic programs. Dropouts were recruited from these schools and organizations providing upgrading programs. Results revealed that graduates perceived themselves as being smarter, more ambitious and responsible, more involved in church and volunteer work, less involved with drugs/alcohol, and less involved in criminal activities than dropouts. Furthermore, graduates were more satisfied with their school experiences and felt parents and teachers were more supportive.

Une approche employant le questionnaire était utilisée pour comparer 92 étudiants finissants du secondaire à 38 étudiants qui ont abandonné leurs études. Les étudiants finissants provenaient d'une école spéciale, d'une école professionnelle et d'une école fondée sur les programmes académiques. Les étudiants qui ont abandonné provenaient d'écoles et d'organisations délivrant des programmes de rattrapage. Les résultats ont révélé que les étudiants finissants se considèrent plus intelligents, plus ambitieux et responsables, plus impliqués à l'église et au travail volontaire, moins reliés aux drogues et à l'alcool et moins impliqués dans des activités criminelles que les étudiants qui ont abandonné leurs études. En outre, les étudiants finissants étaient plus satisfaits avec leurs expériences à l'école et trouvaient plus positifs les parents et les professeurs.

One of the hottest topics in education today is the issue of early school leavers, or as they are commonly referred to, high school dropouts. Over the last seven to 10 years much research has been carried out in the areas of high school dropouts and students who are at risk of dropping out of school (Mann, 1986a; Morris, Pawlovich, & McCall, 1991; Rumberger, 1987; Strother, 1986; Wehlage, Rutter, Smith, Lesko, & Fernandez, 1989). Although the dropout rate of approximately 30% in Canada has remained relatively constant over the last 10 years (Morris et al., 1991), concern about the effects of early school withdrawal has resulted in the highly publicized Canada Employment and Immigration's Stay-In-School Program initiative. This concern has also led to a search for the factors that render students at risk of dropping out prematurely so as to

Dushyanta Persaud teaches mathematics and science at Sisler high school.
Paul Madak is an associate professor in the Department of Educational Psychology. His research interests include school dropouts, sexual abuse prevention, and program evaluation.

identify them and implement prevention/intervention strategies. The purpose of this study was to carry out a direct comparison of individuals who had dropped out of three Winnipeg high schools with individuals who had remained in school until they had graduated, in order to provide educators with additional information that might assist in the reduction/prevention of dropout behavior.

Students who leave school before obtaining a high school diploma place themselves at risk to a lifetime of personal and career problems, as well as creating a problem for Canadian society at large. Dropping out not only reduces the probability that the individual will be able to find full-time employment, but also increases the probability that they will be an added burden to the Canadian tax payer. It has been estimated that unemployed dropouts cost the government billions of dollars in terms of lost tax revenue. Furthermore, there are the added costs associated with the fact that school dropouts are more likely than graduates to need unemployment insurance and/or welfare, be involved with the justice system, and require more health care and other social services (Morris et al., 1991; Rumberger, 1987; Wehlage et al., 1989). The Ontario Ministry of Education reported that in 1981 20% of the dropouts in the 15-24 age group were unemployed, whereas only 1% of the graduates were unemployed (Karp, 1988). Furthermore, a second study in Ontario indicated that the average income of dropouts was approximately $3,770.00 less per year than for graduates (Sullivan, 1988).

The number of individuals who drop out of Canadian schools before obtaining a high school diploma has been reported to vary greatly from province to province and from region to region within provinces. In fact, the dropout rate can vary from a low of 5% to a high of 80% (Morris et al., 1991). This large variation is partly due to the lack of consistent data collection processes across and within the provinces, and partly because the many regional differences cause the rates to vary.

The typical high school dropout has been described in the literature as being predominantly male, from a low socioeconomic and single parent family, having a poor attendance and academic record, above average in age for his or her grade level, not involved in school activities, belonging to a language minority group, and having changed schools a number of times (Beck & Muia, 1980; Hammons, 1988; Mincer, 1989; Morris et al., 1991; Strother, 1986; Wehlage et al., 1989). However, from a review of the research literature on school dropouts, it is clear that no single factor, or group of factors, has been identified as the cause behind an individual's decision to leave school before obtaining a high school diploma. Rather, dropping out of school appears to be a process that occurs over time and is made up of a series of events in the individual's school experience (Finn, 1989; Morris et al., 1991; Rumberger, 1987; Wehlage et al., 1989). Leaving school before obtaining a high school diploma appears to result from a complex interaction between variables and events that begin as soon as the student enters school.

Why Students Drop Out

The reasons for dropping out of school can be broken down into three main categories: (a) social factors; (b) student factors; and, (c) school factors.

Social factors. The social factors that have been found to be associated with dropout behavior include poverty, coming from a single parent family, parents with a low level of education, and belonging to a language minority group (Ekstrom, Goertz, Pollack, & Rock, 1986; Fine, 1986; Johnston, Markle, & Harshbarger, 1986; McDill, Natriello, & Pallas, 1985; Morris et al., 1991; Rumberger, 1983; Steinberg, Blinde, & Chan, 1984; Wehlage & Rutter, 1986; Wehlage et al., 1989). Although it is not clear why individuals with the above characteristics are more likely to drop out of school, it has been hypothesized that the most likely contributing factors include: (a) fewer educational materials found in the home; (b) financial demands requiring older children to hold jobs or staying home to care for younger siblings; (c) parents who have had bad experiences in school and who do not value schooling; (d) parents with lower educational levels providing less assistance with homework; and, (e) language minority students who feel a stronger impact of the above factors because they cannot speak the dominant language (English or French in Canada), which is the instructional language of schools (Ekstrom et al., 1986; Fine, 1986; Rumberger, 1983; Steinberg et al., 1984).

Student factors. Johnston et al. (1986) reported that students who dropped out generally have lower occupational goals for themselves than students who have completed high school. Fine (1986) reported that students who dropped out are generally less motivated to perform in school and feel that although completing high school is important to finding a future job, school did not have anything important to offer to them. Catterall (1987) reported that dropout rates were much higher for those students who had expressed any doubts about finishing school. Furthermore, Johnston et al. (1986), Ekstrom et al. (1986), Wehlage and Rutter (1986), and Wehlage et al. (1989) reported that dropouts take part in fewer after school activities than their counterparts who obtain a high school diploma. This is partly a result of dropouts isolating themselves long before they finally leave school, and partly because a dropout is more likely to have an after school job where he or she is employed for more than 15 hours per week (Fine, 1986; McDill et al., 1985; Rumberger, 1987; Wehlage & Rutter, 1986).

Johnston et al. (1986), Strother (1986) and Kenney (1987) identified pregnancy or child related responsibilities as being one reason why female students drop out of school. However, pregnancy has been identified as a symptom rather than a cause of female dropout behavior. Girls who are performing poorly in school and who lack the basic skills associated with school success were found to be two and a half times more likely to become pregnant than females who were successful at school (Kenney, 1987).

School factors. One of the best predictors of school dropout behavior has been reported to be academic performance. Students with a poor academic history are much more likely to leave school prematurely than are students who are performing at or near grade level. Furthermore, students who dropped out of school were reported to have had poorer work habits and lower locus of control than graduates (Ekstrom et al., 1986; Rumberger, 1983; Wehlage & Rutter, 1986).

Hammack (1986), Lee (1986), Schulz, Toles and Rice (1987) and Shepard and Smith (1989) all reported that students who were overage for their grade

level (i.e., failed and repeated a grade), were more likely to leave school than were their peers who were at the appropriate age/grade level. School attendance has been shown to be another strong predictor of school dropout behavior. Students who eventually dropped out of school have been found to have had sporadic attendance long before they made the final decision to leave school (Broikowske, 1987; McDill et al., 1985; Wehlage & Rutter, 1986; Wehlage et al., 1989).

Thus far this review deals only with factors that are directly related to the student. However, additional factors relate directly to the educational systems where the students attended before dropping out. Researchers such as Fine (1986), Frymier and Gansmeder (1989), Johnston et al. (1986), Wehlage and Rutter (1986), and Wehlage et al. (1989) have reported that students drop out because schools do not offer them the programs nor the caring attention they need if they are to stay in school. School leavers reported that teachers were not really interested in them, nor did teachers understand their problems. For these students school was a place of failure that did not provide the skills relevant for their survival.

Present Research

The research carried out thus far has concentrated on developing profiles of the students who dropped out of school. This type of research provided much needed information as to the general characteristics of those male and female students who leave school early. However, none of the studies that we are aware of has attempted to discover why some students with all the identified general characteristics of a dropout actually drop out of school, whereas others graduate. Therefore, the purpose of this research was to compare dropouts and graduating students who have in common the same general characteristics of a dropout. Specifically, graduates and dropouts were compared regarding their perceptions toward teacher supports, family supports, their academic achievement, involvement in extracurricular activities, study habits, school curriculum, school climate, and overall satisfaction with their educational experience.

Method

Population

Three Winnipeg inner city high schools took take part in this study. These schools were selected because they represented a complete range of educational programing for inner city high school students. School A offered mainly a vocational program, School B offered a mixture of vocational and academic programs, and School C offered mainly an academic program.

School A. This school's enrollment was 275 students at the time the study was carried out and the population was made up of approximately 90% Aboriginal students. The median yearly family income was between $14,000 and $20,000. Sixty two percent of the students were from single parent families. Twenty percent spoke English as a second language. The unemployment rate for the parents of these students was 84% (Winnipeg School Division No. 1, 1990).

School B. Enrollment for this school at the time of the study was 848 students. The population was made up of approximately 60% Aboriginal, 10%

visible minorities (e.g., Portuguese, Filipino, Vietnamese), with the remainder white. The median income for the families of these students was between $14,000 and $20,000. Fifty seven percent of the students were from single parent families, with an unemployment rate of approximately 65%. With regard to the language spoken at home, approximately 11% of the families spoke English as a second language (Winnipeg School Division No. 1, 1990).

School C. The total enrollment for School C at the time of the study was 667 students. School C had the highest income range of the three, varying between approximately $20,001 and $30,000. This school had the lowest unemployment rate at approximately 25%. However, the population at this school had the largest ethnic mix including families of Aboriginal, Portuguese, Filipino, Vietnamese, Italian, Laotian, Chinese, and Polish descent. It also had the highest percentage of English as a second language students (29.1%). Approximately 29% of the students were from single parent families (Winnipeg School Division No. 1, 1990).

Sample recruitment: Graduates. The teaching staff at each of the three schools were asked to identify the students who were to graduate in June 1991, and all potential graduates were approached to take part in the study. For students under 18 years of age, signed parent permission letters were obtained before they were asked to complete the survey. The numbers of graduating students who took part from each of the three schools included seven students from School A, 40 students from School B, and 45 students from School C.

Sample recruitment: Dropouts. For the present study, a dropout was defined as any individual who had left school within the last three years and who would have graduated in June 1991 if they had not left school. However, the identification of the dropout sample did not proceed as expected, because locating dropouts who were willing to participate in the study proved to be difficult. It was hoped that a snowballing procedure using the in-school sample to identify friends who had dropped out would provide the required sample, but this procedure produced only three individuals. The dropouts identified by this method were not willing to complete the survey. Therefore, the majority of the dropouts included were recruited from four local programs designed to provide remedial academic assistance, life skill training, and career exposure to high school dropouts. The dropouts enrolled in these four programs had all left one of the three schools identified above. A total of 35 dropouts who matched the operational definition of a dropout agreed to participate.

Survey Instrument

This study used a 128-item survey that took approximately 55 minutes to complete. The following content areas were included: (a) personal data, which included age, gender, marital status, parental status, employment information, and work habits as a student; (b) school history and attitude toward school; (c) family background data, which included employment status, language spoken at home, parental educational attainment, and socioeconomic status; (d) out of school activities; (e) availability and use of family support structures such as space for doing homework, encouragement for completing high school, and so forth; and, (f) availability and use of school supports such

as counselors, support programs, peer groups, extra assistance from teachers, and so forth.

Some of the items were developed by the authors, whereas other items were taken with permission from Goldfarb's Questionnaire to Graduates (Karp, 1988) and from the Student Assessment Survey (Burstein & Sirotnik, 1984).

This instrument was reviewed by teachers from the three schools and was piloted with four students. Based on the feedback obtained from the teachers and the pilot students, a number of minor changes in question wording were made (a copy of the survey can be obtained from the first author).

Procedure

The study was carried out in May 1990. Parent permission letters were sent home with students who had been identified by their teachers as being most likely to graduate and who were under the age of 18 . For those students over the age of 18, the study was explained to them by the researchers and they were asked to sign a letter stating they agreed to take part in the study. The same procedure was used for the dropouts in the four community programs.

The students still attending school completed the survey during their class time. At School A, the surveys were administered by the researcher who read the questions to them, whereas at Schools B and C the teachers read the questions to the students. For the dropouts, the surveys were administered by the researcher who read the questions to them. The researchers trained the teachers in the administration of the survey in order to ensure that the survey was administered in a standardized manner to all individuals. All data were collected over two weeks.

Results

The data included completed or partly completed surveys from a total of 92 (46 male and 46 female) graduating students and 38 (27 male and 11 female) dropouts. The average age of the graduating students was 18.5, whereas for the dropouts it was 19.5. Of the 92 graduates, 46 (50.0%) spoke only English, whereas 14 (15.2%) also spoke an Aboriginal language, eight (8.7%) Polish, eight (8.7%) French, two (2.2%) Chinese, two (2.2%) Vietnamese, two (2.2%) Ukrainian, two (2.2%) Laotian, two (2.2%) Italian, two (2.2%) Filipino, one (1.1%) Hebrew, one (1.1%) Spanish, one (1.1%) Yugoslavian, and one (1.1%) Hungarian. Of the 38 dropouts, 18 (47.4%) spoke only English whereas eight (21.1%) also spoke an Aboriginal language, five (13.2%) Chinese, two (5.3%) Vietnamese, and three (7.9%) Cantonese.

Because of the high number of chi-square and t-tests carried out, alpha was set at the conservative .01 level in order to reduce the probability of making a Type I error (Hays, 1988).

Social Factors

Employment status. The results indicated that of the individuals who reported the employment status of their parents, 48.9% of the graduates and 23.7% of the dropouts reported that their mothers worked full time. Furthermore, 29.0% of the dropouts and 7.6% of the graduates reported that their mothers collected unemployment or welfare benefits. With regard to their

fathers, 58.7% of the graduates reported them as working full time, compared with 18.4% of fathers of dropouts.

A total of nine (23.7%) dropouts did not know whether their mother was working and 16 (42.1%) did not know if their father was working. For graduates, six (6.5%) did not know their mother's working status, 20 (21.7%) did not know their father's working status.

A total of 54.4% of the 92 graduates reported that they were living in a home owned by their parent(s), whereas 13.2% of the 38 dropouts reported living in a home that was owned by their parent(s).

Family structure. In comparing family structures, it was found that 57.6% of graduates reported living in two parent families whereas 21.1% of the dropouts reported living in two parent families. Of the graduates, 20.7% reported living with a guardian/foster parent, whereas 39.5% of the dropouts were living with a guardian/foster parent.

Language. Approximately equal percentages of dropouts (47.0%) and graduates (50.0%) reported that they could understand at least one language other than English. However, 36.8% of the dropouts and 26.0% of the graduates reported that they lived in homes where a language other than English was predominantly spoken.

Parents' educational level. Of the 92 graduates, 40.2% did not know their father's highest educational attainment, whereas 27.2% did not know their mother's highest educational attainment. For the 38 dropouts who took part in this study, 52.6% did not know their father's educational attainment, whereas 44.7% did not know their mother's educational attainment.

It was found that 30.5% of the fathers of graduates had obtained a grade 12 or higher education level, whereas 13.0% of the fathers of dropouts had done so. Furthermore, 39.2% of the graduates mothers had obtained a grade 12 or higher education level, whereas 15.8% of the mothers of dropouts had done so.

Student Factors

Marital status. A total of 14.1% of the graduates and 10.5% of the dropouts reported being married. Three graduates (3.3%) and three dropouts (7.9%) reported having children with an additional four graduates (4.4%) and one dropout (2.6%) pregnant at the time they took part in this study.

Student employment. Of the 38 dropouts, 46.7% reported that they had worked part time during their last year in school and an additional 21.1% reported having a full-time job just prior to dropping out of school. On the other hand, of the 92 graduates, 21.1% reported working at a part-time job and an additional 3.3% reported working full time.

Self-perceptions. A series of 12 questions asked graduates and dropouts to rate themselves as a student (see Table 1). Based on an analysis of that data, it was found that dropouts were more likely than graduates to rate themselves as being less smart (Chi-square(1)=11.19, p=.001), less ambitious (Chi-square(1)=15.51, p=.001), less helpful to other students (Chi-square(1)=7.4, p=.001), less responsible (Chi-square(1)=15.94, p=.001) and less competitive (Chi-square(1)=6.10, p=.01). Although not reaching statistical significance, it was also found that dropouts were somewhat less self-confident (Chi-

Table 1
Self Rating as a Student

Criteria	Not at all	Only a little bit	Quite a bit	Very much
Hardworking				
Graduates	3.3%	18.5%	51.1%	26.1%
Dropouts	2.6%	31.6%	57.9%	7.9%
Well organized				
Graduates	1.1%	27.2%	57.9%	14.1%
Dropouts	2.6%	31.6%	52.6%	13.2%
Smart				
Graduates	1.1%	13.0%	69.6%	12.0%
Dropouts	7.9%	34.2%	50.0%	7.9%
Friendly				
Graduates	0.0%	7.6%	29.4%	63.0%
Dropouts	0.0%	2.6%	50.0%	60.5%
Self-confident				
Graduates	2.2%	23.9%	48.9%	25.0%
Dropouts	7.9%	34.2%	47.4%	10.5%
Punctual				
Graduates	6.5%	21.7%	44.6%	26.1%
Dropouts	2.6%	31.6%	42.1%	23.7%
Shy				
Graduates	20.7%	41.3%	27.2%	9.8%
Dropouts	10.5%	65.8%	10.5%	13.6%
Ambitious				
Graduates	2.2%	14.1%	48.2%	33.1%
Dropouts	5.3%	44.7%	29.0%	21.1%
Helpful				
Graduates	0.0%	9.8%	43.5%	46.7%
Dropouts	2.6%	26.1%	50.0%	21.1%
Cooperative				
Graduates	0.0%	13.0%	38.0%	46.7%
Dropouts	0.0%	21.1%	42.1%	36.8%
Responsible				
Graduates	1.1%	4.4%	35.9%	57.6%
Dropouts	0.0%	31.6%	42.1%	26.3%
Competitive				
Graduates	9.8%	19.6%	41.3%	28.3%
Dropouts	13.2%	39.5%	31.6%	15.8%

square(1)=3.60, p=.05), and less shy (Chi-square(1)=2.26, p=.13) than were graduates (see Table 1).

Parental expectations and support. A total of 36 questions asked students to rate their parents' support of their educational efforts. Half concerned their perceptions of their father's support, whereas half were concerned with their perceptions of their mother's support. These items dealt with issues such as assistance and encouragement to complete homework, expectation of the effort put forth at school, hopes for the student after high school, encourage-

ment to attend college or university, and expectations to complete high school. To analyze these data, a composite score was developed by adding up the scores for each of the 36 questions. This procedure was used based on Whitehead's (1985) review that concluded that composite scoring was an acceptable method. The results of this procedure produced a mean score for graduates of 84.7 (SD=24.1) and a mean score for dropouts of 74.1 (SD=24.4) out of a total possible score of 144. Using an independence t-test with alpha set at .01, a significant difference between dropouts and graduates was found $t(128)=-2.25$, $p=.01$). Furthermore, when asked where they studied at home, 17.4% of the graduates and 34.2% of the dropouts stated that they "do not study at home." Therefore, these results suggest that graduates perceived that their parents were more supportive of them in their school activities than did dropouts.

School Factors

School attendance. From Table 2 it can be seen that although 42.4% of the 92 graduates had missed fewer than 10 days of school during their last year, 42.1% of the dropouts reported having missed more than 40 days. Furthermore, analysis of the data indicated that the reason most often given by 34.8% of the graduates for missing school was "illness," whereas for 47.4% of the dropouts the most often reported reason was "skipping" school. A total of 22.8% of the graduates gave "skipping" school as their reason for missing school, whereas 29.0% of the dropouts reported that "having a personal problem" was the reason behind most of their absences from school.

Grades repeated. When asked about the number of grades, if any, that they had repeated, 55.4% of the graduates reported that they had not repeated a grade whereas 36.8% of the school leavers reported that they had never repeated a grade. Surprisingly, the same percentage of dropouts and graduates (29%) reported that they had repeated one grade. However, 15.8% of the dropouts had repeated two grades and 15.8% had repeated more than two grades, whereas 5.4% of the graduates had repeated two grades and 6.5% had repeated more than two grades. The differences between graduates and dropouts regarding grades repeated were not found to be statistically different (Chi-square(1)=3.95, $p=.05$).

Courses offered. In order to compare graduate and dropout perceptions concerning the courses offered to them, participants were asked to rate those

Table 2
Attendance During Last Year of School

Attendance	*Graduates*	*Dropouts*
Missed less than 10 days	42.4%	13.2%
Missed between 10-20 days	27.2%	10.5%
Missed between 21-30 days	12.0%	21.0%
Missed between 31-40 days	13.0%	13.2%
Missed more than 41 days	5.4%	42.1%
Totals	100.0%	100.0%

offered to them, or that they had taken. A total of eight questions made up this section of the questionnaire. Question items included, for example, (a) were the courses offered the ones they wanted to take? (b) did the courses teach them what they needed to know? (c) was the difficulty level of the courses appropriate for them? and (d) was resource help available if they needed it? The participants were asked to rate the items on a four point scale from (1) strongly disagree to (4) strongly agree. To analyze these data, a composite score was developed by adding the scores for each of the eight questions. The results of this procedure produced a mean score for graduates of 25.0 (SD=3.2) and a mean score for dropouts of 23.0 (SD=4.0) out of a total possible score of 32. Using an independence t-test with alpha set for .01, a significant difference between dropouts and graduates was found ($t(128)=-3.06$, $p=.005$). Therefore, graduates perceived that the courses offered to them were closer to what they wanted to take than did dropouts.

Teachers as caring. Seven items were utilized to assess what difference, if any, there was between the graduates' and dropouts' perceptions of their teachers. These seven items asked participants to rate teachers on: friendliness, helpfulness, expectations, desire to communicate with them, whether they felt teachers cared about them, and whether they felt teachers did a good job. Compiling composite scores revealed that, out of a total possible score of 28, the mean for the graduates was 22.8 (SD=3.8) whereas the mean for the dropouts was 19.3 (SD=5.8). This difference was found to be statistically significant ($t(128)=-4.1$, $p=.005$). Therefore, graduates were found more likely to perceive their teachers as caring.

Student activities. Two sets of questions asked students about their extracurricular activities. The first set asked participants to indicate in which school activities, if any, they had participated. Table 3 presents the results of those questions. Although it can be seen that some differences do exist between graduates and dropouts, overall these differences were not found to be statistically significant (chi-square(1)= 0.324, $p=.32$). Further analysis of the data indicated that graduates participated in a mean of 2.5 activities, whereas dropouts participated in a mean of 2.3 activities, suggesting that graduates and dropouts participated in the same amount of extracurricular activities.

The second set of questions asked participants to indicate their involvement in a number of community activities. These questions, along with the results, are presented in Table 4. From that table it can be seen that whereas the graduates were more likely to be involved with church (chi-square(1)=8.63, $p<.001$) and volunteer work (chi-square(1)=10.33, $p<.001$), dropouts were more likely to be involved with drug/alcohol use (chi-square(1)=12.86, $p<.001$) and crime (chi-square(1)=10.84, $p<.001$).

Overall satisfaction with their school experience. A total of 12 questions asked graduates and dropouts to rate their overall satisfaction with their school experience. These items asked participants to rate their satisfaction with: the quality of the teachers, types of courses offered, opportunity to learn a career, number of students in their classes, school rules, efforts to encourage completion of school, and so forth. As with previous groupings, a composite score was obtained by adding the 12 items together. Out of a total possible score of 48, the mean score for graduates was 38.0 (SD=5.4) whereas the mean score for

Table 3
Extracurricular Activities Participated in

Activity	Participated in: No	Yes
Sports teams		
Graduates	54.4%	45.7%
Dropouts	39.5%	60.5%
Cheerleading		
Graduates	94.6%	5.4%
Dropouts	97.4%	2.6%
Debating		
Graduates	83.7%	16.3%
Dropouts	92.1%	7.9%
Peer tutor		
Graduates	90.2%	9.8%
Dropouts	84.2%	15.8%
Student council		
Graduates	72.8%	27.2%
Dropouts	89.5%	10.5%
School newspaper		
Graduates	90.2%	9.8%
Dropouts	92.1%	7.9%
Band		
Graduates	90.2%	9.8%
Dropouts	81.6%	18.4%
TV productions		
Graduates	90.2%	9.8%
Dropouts	86.8%	13.2%
Drama		
Graduates	73.9%	26.1%
Dropouts	84.2%	15.8%
Honor assembly		
Graduates	87.0%	13.0%
Dropouts	81.6%	18.4%
Hobby clubs		
Graduates	73.9%	26.1%
Dropouts	79.0%	21.0%
Year book		
Graduates	80.3%	19.6%
Dropouts	68.4%	31.6%
Peer counselling		
Graduates	88.0%	12.0%
Dropouts	94.7%	5.3%
Other		
Graduates	75.0%	25.0%
Dropouts	94.7%	5.3%

Table 4
Community Involvement

Activity	Not involved	Rarely involved	Often involved	Very involved
Church				
Graduates	47.8%	26.0%	15.2%	8.8%
Dropouts	76.3%	21.0%	0.0%	0.0%
Community club				
Graduates	56.5%	21.7%	18.5%	2.2%
Dropouts	39.5%	23.7%	26.3%	10.5%
Cultural activities				
Graduates	53.3%	30.3%	12.0%	4.4%
Dropouts	55.3%	13.2%	15.8%	15.8%
Library				
Graduates	41.3%	34.8%	17.4%	4.4%
Dropouts	63.2%	15.8%	18.4%	2.6%
Volunteer work				
Graduates	33.7%	27.2%	28.3%	10.9%
Dropouts	57.9%	31.6%	10.5%	0.0%
Drug/alcohol use				
Graduates	47.8%	27.2%	19.6%	5.4%
dropouts	18.4%	23.7%	29.0%	29.0%
Crime				
Graduates	89.1%	6.5%	3.3%	0.0%
Dropouts	57.9%	21.1%	13.2%	7.9%

dropouts was 30.3 (SD=7.6). The results of an independent samples t-test found that graduates rated their school experience significantly higher than did dropouts ($t(128)=6.41$, $p=.001$).

Discussion

The purpose of this research was to provide a comparison of students from three inner city high schools who were about to graduate with their peers who had dropped out of school. However, few dropouts were willing to participate in the study, and this led to the procedure of using dropouts enrolled in community programs designed to upgrade their skill levels. Therefore, the results presented here must be viewed in light of the fact that the dropouts who participated in this study might represent a special group. That is, these dropouts, while not attending a formal school, were attempting to upgrade their skill levels. Different results might have been obtained from individuals who had dropped out of school and who were not in the process of upgrading their school related skills.

In addition, the results reported here were obtained from a relatively small number of individuals from the lower income areas of Winnipeg, and neither the schools nor the students were randomly selected. Therefore, (a) the results reported here cannot be generalized to other socioeconomic groups, and (b) while the schools were selected because they were representative of three different types of high school programming, care must be taken when gener-

alizing the results to individuals attending other inner city high schools. Finally, the relatively small sample size resulted in a lack of statistical power, which may have led to Type II errors.

The graduates and dropouts who took part in this study shared many of the same characteristics. That is, graduates as well as dropouts came from areas of an inner city that had high percentages of low income, single parent, and language minority families. Also, both groups were from inner city schools that had high numbers of students who had repeated a grade, high course failure rates, high dropout rates, and students who took part in few in-school extracurricular activities. However, while the results indicated that both groups had similar characteristics (i.e., same socioeconomic status, family backgrounds, etc.), the data mirrored those presented by other researchers that indicated that graduates had fewer of these problems than did peers who left school before obtaining their high school diploma.

Social factors. Although this study was carried out in the inner city section of Winnipeg, the results indicate that graduates were financially better off than were the individuals who had dropped out of school. That is, the parent(s) of graduates were less likely to be unemployed and collecting either social assistance or unemployment benefits, and were more likely to be living in a home that they owned. Furthermore, even though the schools contained high percentages of single parent families, graduates were more likely than dropouts to be living with both parents.

However, although the three schools had high proportions of English as a second language students, language differences were not found between graduates and dropouts. This finding was opposite to that of Steinberg et al. (1984) who reported that dropouts were more likely to be from language minority groups. On the other hand, the education level of parents did seem to be a factor. Graduates reported being more aware of their parents' educational attainments than were dropouts, and further that both their mothers and fathers were more likely to have a grade 12 or higher education levels. The fact that graduates were more aware of their parents' educational attainments suggested that education was more highly valued in those families. This possibility was supported by the fact that graduates reported that their parents were more supportive of their educational efforts and had higher expectations for them than was reported by dropouts.

Student factors. Graduates reported feeling more positive about themselves. They reported feeling smarter, more ambitious, and more helpful to other students than did the individuals who left school before obtaining their high school diploma.

Although graduates repeated fewer grades than did dropouts, this difference was not found to be statistically significant. However, dropouts did report that they skipped school more often than did graduates and also they were more likely to report that they had either a part-time or full-time job at the time they left school.

Although past research suggested that dropouts are less likely to be involved in extracurricular school activities, the results presented above indicate no difference between graduates and dropouts on this variable. However, a difference was found between these two groups for out of school activities.

That is, dropouts were more likely to have reported being involved with drug/alcohol and criminal activities, and less likely to be involved in volunteer and church activities. Furthermore, the fact that we were unable to effectively use the snowballing technique for locating dropouts suggested that graduates and dropouts do not keep in touch with one another once the dropout has left school. It might even be possible that these two groups did not associate with each other even while the dropouts were still attending school. These data suggest that dropouts tend to isolate themselves from their more successful peers (Wehlage, et al., 1989; Wehlage, 1990; Finn, 1989). Further research is needed in order to investigate the pre-dropout relationship between graduates and their peers who leave school early.

School factors. Overall, graduates reported being more satisfied with their school experiences and felt that teachers cared about them. Finally, graduates were also more likely to report that the courses the three schools offered to them were the courses that they wanted to take.

Conclusions

The results presented in this study provide further evidence that school dropout behavior is a complicated issue. Just because a student has repeated a grade, comes from a low income single parent family does not mean that he or she will leave school early. In fact, it was reported by Wehlage (1990) that in the United States "80% of students from single-parents homes do not drop out ... 66% of all dropouts are white; 42% attended suburban high schools; 86% had English as their home language; 71% had never repeated a grade" (p. 2). Those who took part in this study had many characteristics in common, yet some dropped out while others completed their schooling. Wehlage et al. (1989), Wehlage (1990), and Finn (1989) have hypothesized that, when taken together, the evidence concerning why students drop out of school strongly suggests that dropout behavior is a long-term process in which the student slowly becomes disenfranchised from school and the schooling process. Furthermore, these educators believe that the reliance on "'structural characteristics' such as SES or race" (Finn, 1989, p. 133) as the reasons behind dropout behavior is misleading. The model put forward by Finn (1989) called the Participation-Identification Model, suggested that school dropouts have not developed an identification or bonding to schools or the schooling process. Therefore, these students slowly withdraw both emotionally and physically from school. Furthermore, as these students slowly withdraw from school, they take little or no part in school activities either in or out of the classroom, thus increasing the probability that they will not learn to value schooling as an important aspect of their lives. If school does not hold any value, then the final decision to leave without the diploma is easier to make.

Although schools can do little, if anything, to improve the home lives of the students in their charge, they can attend to the fact that dropouts expressed dissatisfaction with their courses, especially with regard to the types of courses offered, course difficulty, and the perceived relevance of the courses to real life work experiences. Schools can develop and offer courses that challenge students and provide them with skills that they perceive as meaningful. Since dropouts also reported that they did not experience warmth and caring from

their teachers, and that they were unhappy generally about their school experiences, schools can attempt to remedy these problems by identifying and addressing the needs of their students.

> But unfortunately, many students report that typical school work is uninteresting. Many at-risk students claim that they do the same things over and over again. Remediation is the curriculum for these students and they see it as boring. Because there is an assumption that "basic skills" must be mastered first, school work fails to challenge students' problem solving skills and therefore lacks qualities that generate intrinsic interest. (Wehlage, 1990, p. 7)

However, as a review of the dropout literature will quickly confirm, there is no single solution to the problem of early school leaving behavior. There is not that magic prevention program which, when adopted by a school, will prevent additional students from dropping out of school. Unless schools are willing to provide a variety of prevention programs designed to reengage students into the learning process, these students will eventually drop out of school. A wide variety of options are needed because not all at-risk students have the same needs, nor do all at-risk students respond to a single set of conditions or curriculum.

> Given the diversity of problems at-risk students encounter, the schools in our study developed strategies in response to these impediments based on two insights: Programs should be matched to student needs and problems, and strategies should take advantage of student interests and strengths. (Wehlage et al., 1989, p. 224)

References

Beck, L., & Muia, J.A. (1980). A portrait of a tragedy: Research findings on the dropout. *High School Journal, 68,* 65-72.

Broikowske, D. (1987). *Truancy: A prelude to dropping out.* Bloomington, IN: National Educational Service.

Burstein, L., & Sirotnik, K.A. (1984). *The use of survey data in comprehensive information systems for local school improvements: Purposes, practices, problems.* UCLA: Center for the Study of Evaluation.

Catterall, J. (1987). An intensive group counseling dropout prevention intervention: Some cautions on isolating at-risk adolescents within high schools. *American Educational Research Journal, 24,* 521-540.

Ekstrom, R.B., Goertz, M.E., Pollack, M., & Rock, D.A. (1986). Who drops out of high school and why? Findings from a national study. *Teachers College Record, 87,* 356-373.

Employment and Immigration Canada. (1990). *Youth: A national stay-in-school initiative.* Ottawa: Author.

Fine, M. (1986). Why urban adolescents drop into and out of public high school. *Teachers College Record, 87,* 393-409.

Finn, J.D. (1989). Withdrawing from school. *Review of Educational Research, 59*(2), 117-142.

Frymier, J., & Gansneder, B. (1989). The Phi Delta Kappa study of students at risk. *Phi Delta Kappan, 71,* 142-146.

Hammack, F.M. (1986). Large school systems' dropout reports: An analysis of definitions, procedures, and findings. *Teachers College Record, 87,* 324-341.

Hammons, R.A. (1988). *Relationships between interschool transfer and dropout rate.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Hays, W.L. (1988). *Statistics* (4th ed.). Toronto, ON: Holt, Rinehart and Winston.

Johnston, J.H., Markle, G.C., & Harshbarger, M. (1986). What research says—About dropouts. *Middle School Journal, 17,* 8-11.

Karp, E. (1988). *The dropout phenomenon in Ontario secondary schools: A report to the Ontario study of the relevance of education and the issue of dropouts.* Toronto: Ontario Ministry of Education.

Kenney, A.M. (1987). Teen pregnancy: An issue for schools. *Phi Delta Kappan, 68,* 728-736.

Lee, J.D. (1986). Education of Native adolescents in inner-city schools. *Canadian Journal of Native Education, 13*(2), 22-26.

Mann, D. (1986a). Can we help dropouts: Thinking about the undoable. *Teachers College Record, 87*, 305-323.

McDill, E.L., Natriello, G., & Pallas, A.M. (1985). Raising standards and retaining students: The impact of the reform recommendations. *Review of Educational Research, 55*, 415-433.

Mincer, J. (1989). Human capital and the labour market: A review of the current research. *Educational Researcher, 18*(4), 27-34.

Morris, S., Pawlovich, W., & McCall, D. (1991). *Evaluating the effectiveness of school drop-out prevention strategies: Some suggestions for future research.* Ottawa: Canadian Education Association.

Rumberger, R.W. (1983). Dropping out of high school: The influence of race, sex, and family background. *American Educational Research Journal, 20*, 199-220.

Rumberger, R.W. (1987). High school dropouts: A review of the issues and evidence. *Review of Educational Research, 57*(2), 101-121.

Schulz, E.M., Toles, R., & Rice, W.K. (1987). The association of dropout rates with student attributes. In W.T. Denton (Ed.), *Dropouts, pushouts, and other casualties* (pp. 89-103). Bloomington, IN: Center on Evaluation, Development and Research: Phi Delta Kappa.

Shepard, L.A., & Smith, M.L. (1989). *Flunking grades: Research and policies on retention.* New York: Falmer.

Steinberg, L., Blinde, P.L., & Chan, K.S. (1984). Dropping out among language minority youth. *Review of Educational Research, 54*, 113-132.

Strother, D.B. (1986). Dropping out. *Phi Delta Kappan, 68*, 325-328.

Sullivan, M. (1988). *A comparative analysis of dropouts and non-dropouts in Ontario secondary schools.* Toronto: Ontario Ministry of Education.

Wehlage, G.G. (1990). The problem of identifying at-risk students. *River East School Division Journal, 2*(1), 1-10.

Wehlage, C.G., & Rutter, R.A. (1986). Dropping out: How much do schools contribute to the problem? *Teachers College Record, 87*, 374-392.

Wehlage, G.G., Rutter, R.A., Smith, G.A., Lesko, N., & Fernandez, R.R. (1989). *Reducing the risk: Schools as communities of support.* Philadelphia, PA: Falmer.

Whitehead, D.J. (1985). *The development and testing of an economics attitude scale.* Research papers in economics Education. London, UK: Institute of Education.

Winnipeg School Division No. 1. (1990). *1988-90 school description summary.* Winnipeg: Author.

For Contributors

Please prepare manuscripts in accordance with these guidelines.

Format

- Manuscript in triplicate.
- Typewritten or printed, double-spaced throughout.
- Abstract of approximately 100 words on a separate sheet.
- Manuscript not to exceed 6,500 words excluding graphics.
- Author's name and affiliation on a separate title page; title only on first page of manuscript to ensure anonymity in the review process.

Referencing

- Sources in parentheses after each reference (direct or otherwise), giving author's name (unless mentioned in text), year of publication, and page number(s).
- Direct quotes of 40 words or less in double quotation marks incorporated into text; those longer than 40 words in indented block format.
- All sources listed alphabetically at the end of the manuscript under the heading *References* using the APA style.
- Explanatory notes, numbered consecutively and marked in the text with superscript numerals, may appear before the References under the heading *Notes;* any citations in notes follow the same format as other references.

Graphics

- Each table and figure on a separate page at the end of the manuscript.
- Figures, charts, and diagrams in camera-ready form.
- Number tables in Arabic numerals and center over each heading, e.g., *Table 1.*
- Number figures in Arabic numerals before the caption at foot of each graph, e.g., *Figure 1: ...*
- Indicate placement of figures and tables in text, i.e., *Insert Table 3 here.*

Style

- *AJER*'s editorial style conforms closely to the *Publication Manual of the American Psychological Association.*
- Editorial changes may be made to manuscripts.
- For spelling, consult *Webster's New Collegiate Dictionary;* spelling in quoted material remains as in the original.

Diskettes

If manuscript is accepted for publication and has been prepared using a Macintosh, IBM, or IBM compatible microcomputer, please send diskette. Microsoft Word, MultiMate, WordPerfect, WordStar, Xerox Writer, XyWrite, and DisplayWrite 3 **only** are acceptable.
Important: Please indicate program used.

Submission to *AJER* implies the manuscript is not being considered for publication elsewhere.

AJER is indexed in the *Alberta Education Index, Canadian Education Index, Current Contents/Social and Behavioral Sciences, ERIC/Current Index to Journals in Education,* and *Social Science Citation Index;* appropriate articles are abstracted in *Cognitive Development Abstracts, Educational Administration Abstracts, Educational Technology Abstracts, Family Abstracts, Language Behavior Abstracts, Multicultural Education Abstracts, Psychological Abstracts, Research into Higher Education Abstracts, School Organization and Management Abstracts,* and *Sociology of Education Abstracts,* and reviewed by the *Journal for Research in Mathematics Education.*

If undelivered return to:

The Alberta Journal of Educational Research
4-116 Education North
University of Alberta
Edmonton, AB
T6G 2G5

Publications Mail Registration Number 1436



THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XXXVIII NUMBER 4 DECEMBER 1992

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

The Alberta Journal of Educational Research

Published in March, June, September, and December
by the Faculty of Education, University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation, and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Editor: Robert H. Short

Consulting Editors

James Chapman
Massey University

Rod Clifton
University of Manitoba

Dianne Common
Lakehead University

Carol Crealock
University of Western Ontario

John Hattie
University of Western Australia

Billie Housego
University of British Columbia

Linda Larocque
University of Alberta

Stewart J.H. McCann
University College of Cape Breton

Les McLean
Ontario Institute for Studies in Education

Linda Phillips
Memorial University of Newfoundland

Sharon Robertson
University of Calgary

Faculty Advisory Committee

John G. Paterson (Chair)
Fern Snart
Joyce Edwards
Graham Fishburne

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada, from the Alberta Advisory Committee for Educational Studies, and from the Office of Research Services, University of Alberta.

The subscription rate is $32.00 per year for individuals, $40.00 per year for institutions. Add $8.00 for delivery outside Canada. Single copies are $8.00 each. Subscriptions and sales in Canada will be charged 7% GST. Please make cheques payable to *The Alberta Journal of Educational Research*. Back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.

Publications Mail Registration Number 1436

Faculty of Education
University of Alberta

The Alberta Journal of Educational Research

Volume XXXVIII, Number 4, December 1992

Sharon G. Neuman	**251**	Editorial: The Negative Consequences of the Self-Esteem Movement
Anne L. Cummings Karen Curtis	**255**	Social Problem Solving Cognitions and Strategies of Student Teachers
Bill Maynes Gordon McIntosh David Mappin	**269**	Experiential Learning Through Computer-Based Simulations
Paul T. Begley Glenda Campbell-Evans	**285**	Socializing Experiences of Aspiring Principals
Barry Onslow Carol Beynon Arthur Geddis	**301**	Developing a Teaching Style: A Dilemma for Student Teachers
Halia C. Silins	**317**	Effective Leadership for School Reform
Tichatonga J. Nhundu	**335**	Job Performance, Role Clarity, and Satisfaction Among Teacher Interns in the Edmonton Public School System

ISSN 0002-4805

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 4, December 1992, 251-253*

Sharon G. Neuman
University of Alberta

Editorial

The Negative Consequences of the Self-Esteem Movement

Editor's Note
In March of this year it was announced that the December 1992 issue of AJER would be devoted to educational research in Quebec. Unfortunately, due to some unforeseen delays with translation, this special issue has been postponed until March 1993. However, please expect a very interesting collection of papers in your next issue of the journal.

The following commentary was submitted by Sharon G. Neuman, a high school English teacher and a graduate student in the area of measurement and evaluation in the Department of Educational Pscyhology at the University of Alberta.

RHS

The self-esteem movement, the attempt in public education to teach or enhance self-esteem at an institutional level, has become a motherhood issue. To question the value of the self-esteem movement is to ally yourself with educational critics and conservatives, which may be why so few questions have been raised from within the educational establishment. Even educators who do have questions may be reluctant to raise them because of their fear of being linked with people they don't agree with in most other ways. However, some questions do need to be asked.

One issue that causes credibility problems for the self-esteem movement is that it appears to be one of the bases for the political correctness movement, whose educational manifestation is censorship. This censorship in the name of never offending anyone and trying to make everyone "feel good" is both intellectually dishonest and educationally unsound. When you reduce everything to bland pap, you are missing great opportunities for discussion and understanding. For example, teaching *The Merchant of Venice* in its historical context provides opportunities for understanding the root causes of anti-semitism and exposing it as the false doctrine it is. Pretending that negative portrayals don't exist doesn't make them go away. The mistake is in stigmatizing words and ideas when we should be dealing with underlying attitudes. One way of doing this is through thorough and open discussion. Certainly our courses of studies should reflect more cultural and ethnic diversity than they

Sharon Neuman is a graduate student in the Department of Educational Psychology at the University of Alberta. She has taught English for many years in public schools. Despite the opinions expressed in this article, she does not consider herself an educational conservative.

have in the past—there is more to literature than dead white males—but they should be inclusive and additive, not exclusive and subtractive.

One of the major questions of the self-esteem movement is (or should be) the source of genuine self-esteem. (I am assuming that there is such a thing, although it hasn't been well defined.) Some educators believe that students must have a positive self-concept before effective learning can take place. But how do students acquire this positive self-concept? Many proponents of the self-esteem movement seem to believe that it is something that can be doled out on the first day of school along with pencils and name tags. Others believe that self-esteem must follow, not precede, real achievement. It is this latter position that makes sense. Self-esteem is important, but no number of self-esteem programs can replace the genuine satisfaction of a job well done. We vastly underestimate our students' understanding and abilities if we think they are deceived by being told how great they are without any kind of evidence. We underestimate their resilience when we attempt to protect them from any kind of failure. We are doing them a great disservice by lying to them with good intentions. Conveying the message that accomplishment isn't important undermines self-esteem by suggesting to students that they lack ability because not much is expected of them.

Schools and school systems often undermine their good intentions in fostering self-esteem by contradicting their own goals. An example of this is encouraging cooperative, noncompetitive learning and then administering and publishing the results of standardized tests. We give our students a mixed message when we suggest that there is no such thing as failure, that their progress will be judged on their own personal growth, and then publicize test results. We are doing our students a disservice by not giving them a safe experience of failure, after which they could go on and succeed, thereby learning that initial difficulties are not insurmountable. Learning to cope with disappointment is very much part of the educational process. We live in a competitive society with a strong emphasis on winning, and children are competitive, as any teacher who has tried to disguise low, medium, and high ability groups with neutral names will confirm.

I suspect that one of the underlying reasons for the popularity of the self-esteem movement in North America is that on the surface we appear to be a concerned, caring, child-centered society, but in fact we're not. Children are not valued in North American culture as they are in many others. The most direct victims of funding cuts in education, health care, and social services are children, perhaps because they can't vote. When you hear stories about some of the excesses of the self-esteem movement, like the principal who installed full-length mirrors with signs saying, "Smile, you're someone special" in school washrooms, you really wonder whether money wouldn't be more productively spent on teaching staff to help students with difficulties. Teaching is difficult and demanding at best, and in the current educational climate where teachers are required to be so much more than transmitters of knowledge it is easy to see why they might be tempted just to say, "That's good" when it really isn't. Education has often been viewed as a panacea for social problems, and as education is required to do more and more outside its traditional bailiwick, giving children self-esteem and solving all these

problems seems an attractive solution. If only it were that easy. Wave your magic self-esteem wand and everyone will be happy. There will be no troubling questions about why Johnny is both illiterate and innumerate. Self-esteem does come through achievement. A great deal of effort and planning must go into programs that genuinely improve self-esteem. Many educators and parents seem not to have the time or the skills to carry these out. It is much easier to give meaningless praise than it is to take the time to show the child how to do a better job.

The self-esteem movement, if continued in its present form, will weaken our educational system, already considered too weak by many of its critics. It will remove from our students any practice in problem solving, as well the joy of accomplishment. If we continue in the present vein, the social and educational problems that the self-esteem movement purports to cure will only become worse.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 4, December 1992, 255-268*

Anne L. Cummings

and

Karen Curtis
University of Western Ontario

Social Problem Solving Cognitions and Strategies of Student Teachers

Social problem solving of student teachers was analyzed using a think-aloud methodology. Participants were asked to solve four problems that student teachers typically face (instruction, discipline, supervising teacher, time demands). Students gave solutions to the problems at the beginning, middle, and end of their preservice year. Problem solutions were analyzed for eight different problem solving elements. The individual strategies generated by the student teachers were content analyzed for each problem situation. Results showed that both the type of problem solving cognitive elements and problem solving strategies were influenced by the type of problem being solved. There were also changes in the problem solving elements over time. Implications of these results for teacher education are discussed.

On analysa la résolution des problèmes sociaux des stagiaires en utilisant une méthodologie de réflection à haute voix. Les participant(e)s ont dû résoudre quatre problèmes typiques que les stagiaires rencontrent tels la pédagogie, la gestion de classe, l'enseignant(e) coopérant(e), l'exigence de temps. Les stagiaires ont solutionnés ces problèmes au début, au milieu et à la fin de leur année dans laquelle ils (elles) complétaient leur dernier stage. Les solutions aux problèmes ont été analysées selon huit critères différents et le contenu de chaque situation problématique a été analysé d'après les stratégies individuelles utilisées par les stagiaires. Les résultats obtenus indiquent que la nature même du problème étant résolue influencent le choix d'éléments cognitifs et les stratégies de résolutions préférées par le stagiaire. Il y eut aussi des modifications à la résolution de problèmes pendant la période de temps étudiée. Les implications que ces résultats apportent à la formation des enseignant(e)s sont discutées.

Educational research has begun to focus increasingly on teacher thinking and problem solving (Cummings, Murray, & Martin, 1989; Kagan, 1988; Peterson, 1988). This focus represents a significant shift away from more behavioral research. As Clark and Lampert (1986) assert, teacher education that is based only on behavioral research often leads to simplified prescriptions for teacher performance (e.g., how to reinforce appropriate student behaviors). In contrast, research on teacher thinking views the thought processes, judgments, and decisions of teachers as having a profound impact on the way teachers teach and students learn in classrooms (Peterson, 1988).

Research on teacher thinking has the potential of contributing greatly to teacher education programs. Flodden and Klinzing (1990) suggest that insight

Anne L. Cummings is an assistant professor of counseling psychology. Her research interests include social problem solving, counseling process, and women's issues in counseling.
Karen Curtis recently completed her MA in school psychology from the Ontario Institute for Studies in Education.

into teachers' thought processes could provide information for both content and methods of teacher education. For example, if teacher educators knew the schemata that teachers employed to help them deal with various teaching situations, it might then be possible to facilitate the development of these schemata in student teachers.

Clark (1988) emphasizes the importance of research on teacher thinking as being inherently suited to the complexity and uncertainty that is intrinsic to the interactive teaching experience. He believes that much of a teacher's energy is directed to predicting problems and reacting to the interruptions and dilemmas that they encounter daily. In fact he reports that teachers are faced with interactive decision situations at approximately two-minute intervals while teaching. If we can better understand this interactive decision making process used by teachers, teacher educators may be better able to help student teachers prepare themselves to think in ways that will allow them to cope with the complexities of the classroom. By viewing teachers as thoughtful, active problem solvers, an external perspective of process-product research (e.g., focusing on what teachers do) is then replaced by an internal perspective of what teachers understand and perceive (Kagan, 1988).

While much research has examined well-defined problem solving in the classroom using math or science problems (e.g., Leinhardt & Smith, 1985; Peterson, 1988), little research has investigated the problem solving of teachers with more ill-defined social problems that arise daily in the classroom. Social or personal problem solving is here defined as the process utilized by people to cope effectively with the problematic situations of everyday life (D'Zurilla & Goldfried, 1971). It is conceptualized as being highly interactive, complex, and intermittent rather than being a linear or stage sequence (Heppner & Krauskopf, 1987).

Several researchers (Berliner, 1986; Peterson & Comeaux, 1987) note that beginning teachers, as opposed to more experienced teachers, tend to be at a disadvantage in problem situations because of their lack of both procedural and declarative knowledge of common teaching practices. They do not yet have a broad repertoire of hypotheses or schemata for responding to the constantly changing classroom environment. Flodden and Klinzing (1990) recommend studying the implicit theories of student teachers to facilitate the development of schemata to achieve a wide range of educational aims. Therefore, the present study will focus on the social problem solving of student teachers.

One difficulty in researching teacher thinking is that thinking cannot be accessed directly. One of the most effective ways of dealing with this problem is to assess thinking and problem solving processes through think-aloud methodology (Clark, 1988; Ericsson & Simon, 1984). This methodology requires teachers to talk out loud while they solve problems, giving researchers a close approximation of the cognitive and metacognitive strategies employed by teachers as they reflect on their own actions and knowledge about a given situation.

The think-aloud methodology has been used by one research team (Magliaro, Wildman, Niles, McLaughlin, & Ferro, 1989; Wildman, Magliaro, Niles, McLaughlin, & Drill, 1988) to identify the specific problems en-

countered by elementary and secondary teachers during their first two years of teaching. These researchers found several primary sources of problems for beginning teachers: students, instruction, classroom organization, school context, self, and parents. Although the Wildman et al. (1988) study focused on beginning teachers, the present study focuses on student teachers. Four of the areas identified by Wildman et al. (1988) were selected for examination in the present study as being representative of the types of ill-defined problems that all student teachers experience: instruction, classroom organization (discipline), context (relationships with colleagues), and self.

The present study was designed to accomplish several goals. The first was to analyze the think-aloud problem solutions of student teachers on four different types of problems to identify the cognitive elements used and to assess whether the students varied their cognitive processing for different types of problems. Eight problem solving elements identified by Cummings et al. (1989) were used for the analyses. One of the weaknesses of this previous study was the use of written problem solutions rather than oral solutions that might permit greater completeness of the thoughts expressed.

The second purpose involved assessing the relationship between different types of problems and the types of problem solving strategies used by student teachers. Previous research (Cummings et al., 1989; Kagan, 1988) has found that, to a large extent, the type of social problem encountered influences the kind of problem solving approach. However, previous studies have not examined the content of the strategies selected by participants. Understanding how different types of problems might elicit different strategies from student teachers could aid teacher education in better preparing students for typical classroom situations. For example, do student teachers use different types of strategies for discipline problems than they do for instruction problems?

Finally, the effect of experience and knowledge on the problem solving processes of the student teachers was examined. This goal was accomplished by assessing students' problem solving processes at the beginning of their student teacher education before they had any experience in the classroom, in the middle of their training after some education in procedural knowledge (practice teaching) and declarative knowledge (university education courses), and toward the end of their teacher preparation year.

One of the weaknesses of previous studies on teacher cognitions using a think-aloud methodology is that sample sizes rarely exceed 10 (Kagan, 1990), making generalizability of results difficult. Two studies that did use larger samples (Stoiber, 1991; Swanson, O'Connor, & Cooney, 1990) only examined solutions for discipline problems. Therefore, the present study attempts to rectify this problem by using a larger sample of 29 student teachers solving four diverse social problems.

Method

Participants

The participants were 29 (22 women and 7 men) volunteer student teachers in a one-year Bachelor of Education program at the University of Western Ontario during the 1988-1989 academic year. All students had a three- or four-year bachelor's degree before entering the Bachelor of Education program. Ages ranged from 23-43 with a mean of 27. Students were preparing to teach

grades 4-12 and had a wide variety of majors (e.g., English, French, history, music, math). The organization of the program included a repeated sequence throughout the year of students taking five weeks of course work at the university followed by three to four weeks of practice teaching in schools. During each practice teaching session students were assigned to one supervisory teacher and taught portions of that teacher's load. The students' placements changed for each practice teaching segment.

Procedure

A research assistant met individually with the volunteer student teachers during the first weeks of the academic year in September. The following instructions were read to each student: "You will be given four different problems. For each one of them, put yourself in the situation and think aloud about how you would resolve the problem to a satisfactory solution. I am interested in your thoughts, feelings, and behaviors. Read the situation out loud and then begin sharing your thoughts about the problem. Do not censor anything." All problem solutions were audiotaped as the student spoke out loud.

The four problem situations included the following: (a) "You have just taught a new concept to the class. The brighter students understand the new concept, while the slower students do not understand it" (instruction); (b) "Two students together have been gradually escalating their disruption of your class presentations over time and have not responded to your requests to be quiet" (discipline); (c) "You are having conflict with your supervising teacher because of differences in style and approach" (context); (d) "You are finding it difficult to balance the demands of school work and the needs of your personal life" (self).

These four problem situations were repeated twice more for each student teacher. The second testing occurred two months later in November after the first practice teaching experience of four weeks. The last testing occurred six months later in March after the second practice teaching session of three weeks. In both cases the research assistant met with the student teachers individually and taped their think-aloud solutions to the same four problems.

All problem solutions were transcribed from the audiotapes. After the transcriptions were completed, the same research assistant listened to the tapes and made any necessary corrections to the typed transcripts. Two raters then divided the problem solutions into scorable units and categorized the units using a manual developed by Cummings et al. (1989). The scoring manual specified procedures for unitizing scorable units of analysis and for assigning these units to categories. A scorable unit is denoted by a complete thought or change of content within a sentence.

Protocols were scored for nine mutually exclusive categories intended to describe the total process of social problem solving. The scoring categories were based on problem solving elements from prominent social problem solving models (D'Zurilla & Goldfried, 1971; Platt & Spivack, 1975). The nine categories were: (a) problem statement (describing the problem or "setting the stage" before taking action); (b) analysis (directly or indirectly assessing the problem); (c) strategy (idea, action, or approach for solving the problem); (d) elaboration (detail about a strategy or the situation); (e) reason (explanation for

a strategy, feeling, or evaluation); (f) evaluation of a strategy or situation (judging the feasibility or desirability of a strategy or situation); (g) evaluation of self (problem solver judges self within the problem situation); (h) feeling (how the problem solver feels about the problem, strategy, or outcome); (i) if-then thinking (assumption about the problem, followed by a possible solution). Because of the low frequency of the category *problem statements* by these students, the first category above was combined with elaboration, yielding eight problem solving elements. The combination of problem statement with elaboration was conceptually sound in that both categories could be seen as descriptive detail about the problem. An example from one student teacher on the discipline problem of the unitizing and coding categories includes the following: "I might move them closer to me (strategy) so I could have more eye contact with them (reason). The only risk in that is possible embarrassment at being in front of their peers (evaluation of strategy)."

A frequency count for each element for each of the four problems was made from the transcribed problem solutions. The frequencies were then totalled for the two conflict problems (discipline, supervising teacher) and the two balancing problems (instruction, time demands). The conflict problems involved solving an interpersonal conflict between the student teacher and other people (students in one problem and supervising teacher in the other). Balancing problems involved two opposing needs or pressures (balancing the needs of the slow versus fast students in the instruction problem, and balancing the needs of work versus home life in the self problem). A second grouping occurred by combining the frequencies for the two classroom problems (discipline, instruction) and the two nonclassroom problems (supervising teacher, time demands). These frequencies were used in subsequent analyses.

Protocols were scored by two raters: one rater had an MEd degree and five years of experience in rating protocols, while the other rater was the same research assistant previously mentioned, a graduate student in an MEd program. Raters received approximately 20 hours of training and practice on unitizing and categorizing practice protocols produced by people who did not participate in the study. Training was discontinued when raters reached an acceptable level of interrater reliability. Interrater agreement on dividing the protocols into scorable units was 90%. Cohen's (1960) coefficient of interjudge agreement, calculated at the end of training for the nine problem solving categories, was .88.

One of the weaknesses of the Cummings et al. (1989) study was that the content of the problem solutions was not analyzed for the types of strategies employed. To address this issue, content analysis of the strategies generated for the four problems was performed. The rating was done by two different raters from the research assistants: the authors. Using protocol analysis techniques (Ericsson & Simon, 1984), seven strategy categories were developed for three of the problems and six categories were developed for the fourth problem. As can be seen in Table 1, the instruction and discipline problems share four common categories, whereas the supervising teacher and time demands problem also share four common categories. Cohen's (1960) coefficient of interrater agreement, calculated for all of the strategy categories, was .87.

Table 1
Problem-Solving Strategy Categories

Instruction Problem	*Discipline Problem*
1. try a different approach	1. try a different approach
2. assess the situation	2. assess the situation
3. deal with the problem out of class	3. deal with the problem out of class
4. seek help from others	4. seek help from others
5. repeat or review	5. talk about the problem
6. deal with the problem in class on an individual basis	6. take direct action
7. deal with the problem in class with an emphasis on small groups	7. use an indirect solution
Supervising Teacher Problem	*Time Demands Problem*
1. talk about the problem	1. talk about the problem
2. use self-talk	2. use self-talk
3. assess the situation	3. assess the situation
4. compromise/find a balance	4. compromise/find a balance
5. seek help from others	5. organize/establish priorities
6. use an indirect solution	6. take direct action
7. look for the positive	

Frequency counts were made for each category for each of the four problems to be used in the analyses.

Results

The first goal of the present study was to assess whether student teachers varied their cognitive processing for different types of problems. To address this goal, the four problem situations were grouped in two different ways (conflict vs. balance; classroom vs. nonclassroom) and two MANOVAs were performed.

In the first MANOVA, problem situations were grouped into conflict problems (discipline, supervisor) and balance problems (instruction and time demands). Independent factors were two levels of problems (conflict and balance) and three levels of testing (early, middle, late). Frequencies of occurrence of the eight problem solving elements were the dependent measures. There was a statistically reliable main effect of problem type on the frequency scores for the problem solving elements considered collectively, multivariate $F(8,21)=4.19$, $p=.004$. This effect was traced to two statistically reliable univariate F values for two of the eight problem solving elements. On interpersonal, conflict problems, participants used more evaluation of strategies, $F(1,28) = 5.08$, $p=.03$, and provided more reasons, $F(1,28) = 22.76$, $p<.001$. Thus students were more analytical and evaluative of their actions when the problem involved other people (see Figure 1).

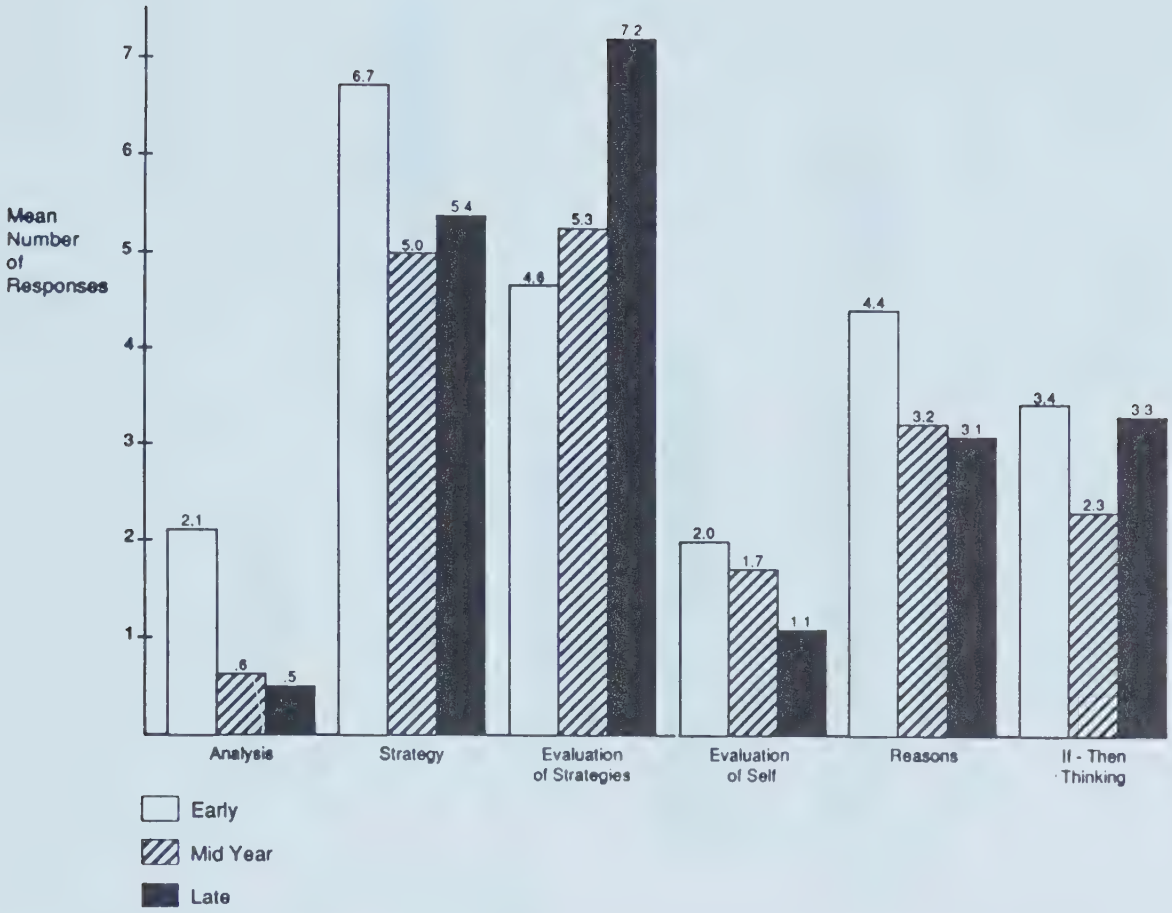


Figure 1. Significant differences in mean number of problem solving elements on balance versus conflict problems.

When the problems were combined into two classroom problems (instruction, discipline) and two nonclassroom problems (supervisor, time demands) for a second MANOVA (with three levels of time of testing again constituting a second independent factor), there was a statistically reliable main effect for problem type, $F(8,21) = 7.88$, $p<.001$. In this case, there were statistically reliable univariate F values for four of the eight problem solving elements. In solving nonclassroom problems, participants used more elaboration, $F(1.28)=6.51$, $p=.02$; more evaluation of strategies, $F(1,28)=10.04$, $p=.004$; and more evaluation of self, $F(1,28)=43.04$, $p<.001$. However, on classroom problems, participants generated more strategies, $F(1,28)=35.50$, $p<.001$. Mean values for the problem solving elements highlighted in these analyses are displayed in Figure 2.

The second goal of the study was to determine the relationship between different types of problems and the type of problem solving strategies used by student teachers. To determine whether there were statistically reliable differences in category usage, strategy categories that were common to two problems were compared using three MANOVAs. For the first one-way MANOVA, the independent variable was two levels of classroom problems (discipline and instruction) and the dependent variables four (of seven) common strategy categories for those two problems (try a different approach, assess the situation, deal with problem outside of class, and seek help from

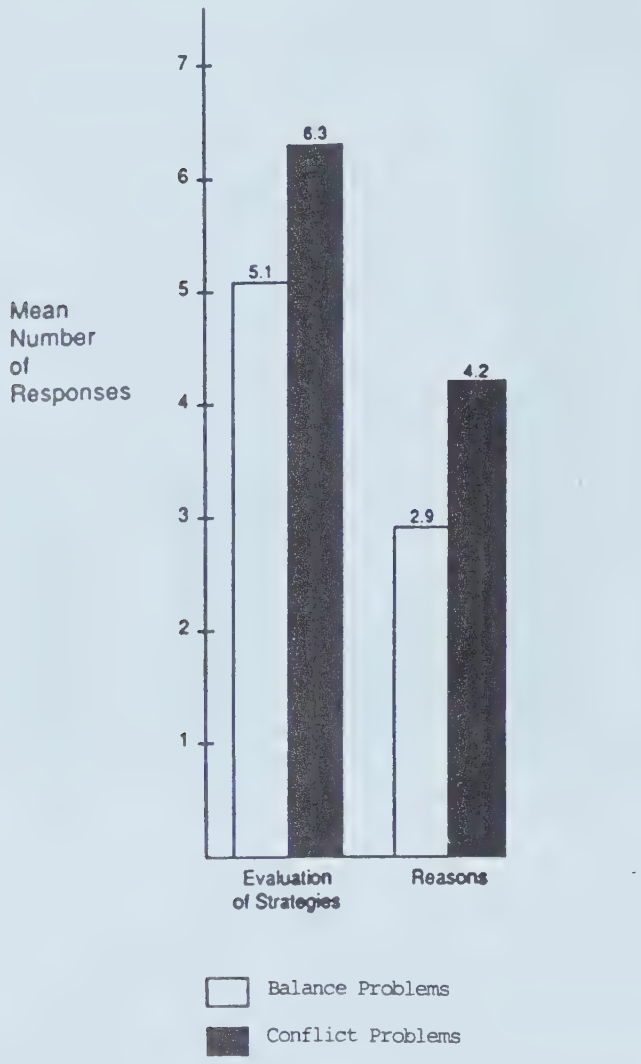


Figure 2. Significant differences in mean number of problem solving elements on classroom versus nonclassroom problems.

others). There was a statistically reliable main effect of problem on the frequency of scores for the strategy categories, $F(4,24)=7.28$, $p=.001$. This effect appeared to be due to two statistically reliable univariate f values for "try a different approach," $F(1,27)=11.85$, $p=.002$, and "assess the situation," $F(1,27)=6.72$, $p=.015$. Participants used both of these strategies more often on the instruction problem than they did on the discipline problem.

A second one-way MANOVA compared the two nonclassroom problems (supervising teacher and time demands) on their four common strategies (talk about the problem, use self-talk, assess situation, and compromise/find a balance). Once again, there was a statistically reliable main effect for problem, $F(4,24)=6.69$, $p=.001$. This effect appears to be due to two statistically significant univariate F values for "talk about the problem," $F(1,27)=25.75$, $p=.001$, and "assess the situation," $F(1,27)=5.02$, $p=.03$. Students used "talk about the problem" more often for the supervising teacher and "assess the situation" more often for dealing with time demands.

A final one-way MANOVA compared the four common categories for the interpersonal conflict problems (discipline and supervising teacher) on their four common strategies (assess situation, seek help from others, talk about the problem, and use an indirect approach). There was a statistically reliable main

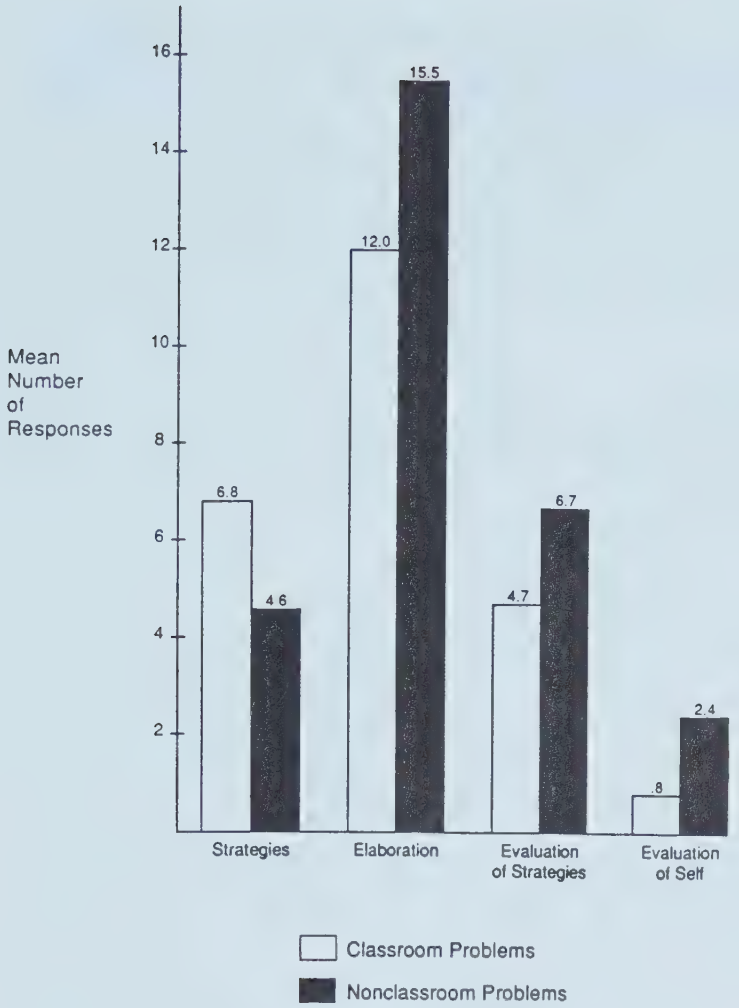


Figure 3. Significant differences in mean number of problem solving elements at three testing times: pre, mid-year, and post.

effect for problem, $F(4,24)=3.29$, $p=.03$. Two statistically significant univariate F values include, "assess the situation," $F(1,27)=6.67$, $p=.015$, and "use an indirect solution," $F(1,27)=5.86$, $p=.02$. Participants assessed more often on the discipline problem and used an indirect solution more often when dealing with the supervising teacher.

The third purpose of the study focused on examining student teachers' problem solving over time to determine any effect of experience and knowledge gained during teacher education. There was a statistically reliable main effect for time on the problem solving elements when all four problems were combined, $F(16,94)=2.20$, $p=.01$. This effect appeared to be traceable to three statistically significant univariate F values for analysis, $F(2,26)=4.08$, $p=.03$, strategies, $F(2,26)=3.55$, $p=.04$, and evaluation of strategies, $F(2,26)=3.62$, $p=.04$. Over the course of the school year, these student teachers used less analysis, fewer strategies, and more evaluation of strategies as shown in Figure 3.

On the surface, the decrease in strategies and analysis over time is puzzling. One would expect that the experience and knowledge gained during teacher education would lead to more analysis and a wider range of strategies rather than to fewer of each. To gain clarity about these results, it became necessary to content analyze the problem responses for each participant for each problem over the three testings. Although there were many individual differences in the responses, several patterns emerged from this analysis. We

need to emphasize that the following patterns can only be seen as possible trends in the data, as they were not able to be assessed through statistical tests.

Through reading the problem solutions, it became clear why there was a decrease in strategies and analysis over time. For example, on the first testing for the discipline problem, many participants would list a number of strategies that they might try. At this point in their program, the preservice teachers had not experienced a practice teaching assignment and, thus many of their problem solutions were speculative about what might work. For example, in response to the discipline problem, one participant said, "I would go to that part of the room where the students are being disruptive. If that doesn't work, I would go back again and ask them to please be quiet. If that didn't work again, I would go and sit behind them and lecture right behind them. If they were still being disruptive, I would start directing questions at them so they would have to be on their toes and get them involved...."

At the second testing after a practice teaching period of four weeks in the schools, the same participant had a better idea of what would work. "I had some general level classes and I had a lot of kids who were disruptive. I tried to break up the class into three different teaching styles. We had 75-minute periods so I had a lesson, a lecture, and I moved into individual work where a student had to hand in an assignment afterwards, and then I moved to group work. I was hopefully trying to keep them busy to avoid the situation." In this case, the preservice teacher did not have to consider many different strategies, as in the first answer, because she had found a strategy that worked for her from her experience.

Discussion

With the complexities inherent in the modern classroom, it is becoming essential that teacher education and educational research be concerned with the thinking and problem solving processes of teachers. We need to understand better the thought processes of teachers as they traverse the complex matrix of ill-defined classroom problems. By examining in detail the thought processes of student teachers as they solve typical problems in teachers' lives, we can increase our understanding of the natural problem solving processes that student teachers use and begin to determine how best to supplement these natural processes.

As in the Cummings et al. (1989) study, the type of problem influenced both the type of problem solving elements and the type of strategies employed by these student teachers. Problems of interpersonal conflict elicited significantly more evaluation of strategies and reasons, perhaps indicating that the students were more reflective about their actions when conflict was involved.

When problems were divided into classroom and nonclassroom problems, student teachers were able to generate significantly more strategies for the classroom situations. This result is not surprising given that much of their teacher education focused on specific ways of handling various classroom dilemmas.

With nonclassroom problems, students used significantly more elaboration (descriptive detail) and more evaluation, both of themselves and of the strategies they were employing. In particular, it is interesting that students

used more evaluation for nonclassroom problems than for classroom problems. When this result is combined with the greater number of strategies for classroom problems, it may indicate that these students were inclined to be more action-oriented and less reflective about their actions when dealing with classroom situations. In fact, research on problem solving consistently finds that compared with experts, novices do not reflect on their problem solving and work more on the surface features of problems (Chi, Glaser, & Farr, 1988; Eylon & Linn, 1988; Glaser, 1990).

As expected, there were changes in students' usage of the problem solving elements over time. On all four problems combined, there was a decrease in analysis and number of strategies and an increase in evaluation of strategies. It appears from the content analysis that the decrease in analysis and strategies may represent student teachers becoming clearer as a result of experience about what strategies are effective for different situations. They therefore may not need to analyze the situation as much and could be more parsimonious in their choice of strategies. In other words, they learn "what works." This interpretation is consistent with research on expert-novice differences in problem solving (e.g., Chi et al., 1988), which shows that experts more than novices activate a parsimonious set of schemata to solve problems efficiently. Of course, it is also possible that students experienced a practice effect from responding to the same problems over a six-month period that led to fewer strategies being generated.

The increase in evaluation of strategies over the course of the year is a promising finding in that one goal of teacher education programs is to help students learn to think reflectively so that they can deal with the complex problems of the classroom. It is a hopeful sign that these students were evaluating their strategies more at the end of the year than they were at the beginning. Unfortunately, we do not know if this change is a result of theoretical knowledge or practice teaching experience or some combination of the two.

When the content of the strategies chosen by students was examined, interesting patterns emerged. From the different comparisons between problems, the strategy "assessing the situation" was used significantly more often on three of the four problems. The problem that had the least frequency of this strategy involved conflict with a supervising teacher. When this result is combined with the finding that students also used an indirect solution more often with the supervising teacher problem than with the discipline problem, a possible interpretation emerges. Perhaps when students are in a position of less power and control, they are less likely to reflect or assess and more likely to take an indirect approach. In other words, indirect solutions may be elicited by the unequal power in the relationship between student teacher and supervising teacher. This hypothesis is further strengthened by considering that when the participants were in a position of greater power relative to pupils on the discipline problem, the most common strategy was "take direct action."

Another, more general, observation arises from examination of Table 1, which shows that similar problems (e.g., classroom problems of discipline and instruction) elicited both some similar strategies and some different strategies. On the other hand problems that were quite different (e.g., time demands and

instruction) elicited few similar strategies. Thus, although there are some commonalities in approaches to similar problems, the more disparate problems become, the less commonality there appears to be in efforts to resolve them.

Traditional problem solving approaches (D'Zurilla & Goldfried, 1971; Platt & Spivack, 1975), however, have suggested a rather set, linear approach to all problem solving, viewing successful problem solving as a progression through a series of previously defined steps regardless of the type of problem. The results of the present study suggest that problem solving solutions are not generalizable across persons and contexts. Perhaps preset problem solving approaches are better suited to well-defined problems of math and science than to ill-defined social problems such as those investigated in this study.

The results of this study provide useful information about student teachers' thought processes while solving typical problems. In accordance with previous research (Cummings et al., 1989), the results seem to suggest that the type of problem influences the type of cognitive processing and strategies used. As Kagan (1988) asserts, the kinds of interactive decisions made by teachers and how they shape practical problem solving are determined by the specific task environments encountered.

To address these challenges, teacher education programs might best meet the needs of student teachers by facilitating the development of context-specific schemata as well as providing general strategies that would apply across similar situations. In reviewing studies on transfer of learning, Perkins and Salomon (1989) found that effective thinking was quite dependent on context-bound knowledge and skills that do not easily transfer to other domains, thus supporting the need for cognitive training that is rooted in specific content. Brown, Collins, and Duguid (1989) refer to the situated nature of knowledge and recommend the learning of such schemata through a cognitive apprenticeship.

However, Perkins and Solomon (1989) also discovered in more recent research that transfer can occur when general principles of reasoning are taught in conjunction with self-monitoring activities and potential applications to other possible contexts, thus accenting the need for more general schemata. As Swanson et al. (1990) note, if pattern recognition is important in fields such as physics and medicine, chances are high that it will be helpful for classroom problems also. Therefore, instead of teacher education being focused on delivering facts and rules, the emphasis needs to shift to joining context-specific knowledge for solving ill-defined human problems with more general strategic knowledge.

To achieve transfer of knowledge in ill-structured domains, some authors (Spiro, Vispoel, Schmitz, Samarapungavan, & Boerger, 1987) contend that multiple interconnections among different aspects of domain knowledge need to occur in conjunction with multidimensional representations of many examples. They believe that there are no rules general enough to cover most cases. What is required, instead, is a flexibility of thinking in how knowledge is represented in memory.

Teacher education programs, then, need to encourage flexibility of student thinking. One such project is described by Stoiber (1991). Her reflective in-

struction program was designed to help preservice teachers become more aware of their decision making abilities and the differing cognitive capabilities required by different problem situations. She found that having preservice teachers think aloud about their classroom problems during teacher education experiences enhanced their cognitive monitoring and evaluative functioning.

It must be noted that the generalizability of the results of the present study is limited by the unequal sex ratio of the sample, the wide variety of grades and disciplines they were preparing to teach, and the type of teaching practicum program experienced by these student teachers. The study is further delimited by the description of classroom scenarios with only two sentences, which does not capture the complexity of classroom situations. These more narrow definitions of problems may constrict the applicability of the findings because participants' responses may not have been representative of their actual behavior in the classroom. Future research needs to study problem solving of more complex, ill-defined classroom problems through the observation of student teachers as they actively solve problems in a natural setting. In addition, having teachers think aloud about actual problem situations in classrooms through the use of stimulated recall techniques could provide helpful information about more in vivo thought processes.

Acknowledgments

Preliminary versions of sections of this article were presented at the Annual Meeting of the American Educational Research Association, Boston, MA, April 1990 and the Canadian Psychological Association Convention, Ottawa, Ontario, June 1990. We gratefully acknowledge the assistance of Gillian Villanueva and Rita Prettie in the data collection and coding process and Jack Martin for comments on earlier drafts of this article. Correspondence regarding this article should be sent to Dr. Anne Cummings, Faculty of Education, The University of Western Ontario, London, Ontario N6G 1G7.

References

Berliner, D.C. (1986). In pursuit of the expert pedagogue. *Educational Researcher, 15*(7), 5-13.

Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Chi, M., Glaser, R., & Farr, M.J. (1988). *The nature of expertise.* Hillsdale, NJ: Erlbaum.

Clark, C.M. (1988). Asking the right questions about teacher preparation: Contributions of research on teacher thinking. *Educational Researcher, 17*(2), 5-12.

Clark, C. M., & Lampert, M. (1986). The study of teacher thinking: Implications for teacher education. *Journal of Teacher Education, 37*(5), 27-31.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurements, 20*, 37-46.

Cummings, A.L., Murray, H.G., & Martin, J. (1989). Protocol analysis of the social problem solving of teachers. *American Educational Research Journal, 26*, 25-43.

D'Zurilla, T.J., & Goldfried, M.R. (1971). Problem solving and behavior modification. *Journal of Abnormal Psychology, 78*, 107-126.

Ericsson, K.A., & Simon, H.A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Eylon, B.S., & Linn, M.C. (1988). Learning and instruction: An examination of four research perspectives in science education. *Review of Educational Research, 58*, 251-301.

Flodden, R.E., & Klinzing, H.G. (1990). What can research on teacher thinking contribute to teacher preparation? A second opinion. *Educational Researcher, 19*(4), 15-20.

Glaser, R. (1990). The reemergence of learning theory within instructional research. *American Psychologist, 45*, 29-39.

Heppner, P.P., & Krauskopf, C.J. (1987). An information processing approach to personal problem solving. *The Counseling Psychologist, 15*, 371-447.

Kagan, D.M. (1988). Teaching as clinical problem solving: A critical examination of the analogy and its implications. *Review of Educational Research, 58*, 482-505.

Kagan, D.M. (1990). Teacher cognition. *Review of Educational Research, 60*, 419-469.
Leinhardt, G., & Smith, D. (1985). Expertise in mathematics instruction: Subject matter knowledge. *Journal of Educational Psychology, 77*, 241-247.
Magliaro, S.G., Wildman, T.M., Niles, J.A., McLaughlin, R.A., & Ferro, S. (1989, March). *Modeling the transition from novice to expert: A three-year study of teacher problem solving*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
Perkins, D.N., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher, 18*(1), 16-25.
Peterson, P.L. (1988). Teachers' and students' cognitional knowledge for classroom teaching and learning. *Educational Researcher, 17*(5), 5-14.
Peterson, P.L., & Comeaux, M.A. (1987). Teachers' schemata for classroom events: The mental scaffolding of teachers' thinking during classroom instruction. *Teaching and Teacher Education, 3*, 319-331.
Platt, J.J., & Spivack, G. (1975). *Manual for the means-ends problem-solving procedure (MEPS): A measure of interpersonal cognitive problem-solving skill*. Philadelphia, PA: Hahnemann Community Mental Health/Mental Retardation Center.
Spiro, R.J., Vispoel, W.P., Schmitz, J.G., Samarapungavan, A., & Boerger, A.E. (1987). Knowledge of acquisition for application: Cognitive flexibility and transfer in complex content domains. In B.C. Britton (Ed.), *Executive control processes* (pp. 177-199). Hillsdale, NJ: Erlbaum.
Stoiber, K.C. (1991). The effect of technical and reflective preservice instruction on pedagogical reasoning and problem solving. *Journal of Teacher Education, 42*, 131-139.
Swanson, H.L., O'Connor, J.E., & Cooney, J.B. (1990). An information processing analysis of expert and novice teachers' problem solving. *American Educational Research Journal, 27*, 532-556.
Wildman, T.M., Magliaro, S.G., Niles, J.A., McLaughlin, R.A., & Drill, L. (1988, April). *Sources of teaching problems and the ways beginners solve them: An analysis of the first two years*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 4, December 1992, 269-284*

Bill Maynes

Gordon McIntosh

and

David Mappin
University of Alberta

Experiential Learning Through Computer-Based Simulations

During the 1980s, many university-based preparation programs for prospective school administrators introduced practicum- or internship-like courses in efforts to establish a better balance between theory- and experience-based learnings for their students. This article reports on a project that is investigating the potential of a sophisticated interactive laser videodisc simulation to provide an alternative source of experiential learning opportunities. The major question asked is, In an experiential learning cycle, to what extent can the simulation effectively replace the real work of a principal? The answer is sought through an examination of the experiences of two distinct subgroups of students: a small number who seem to have gained very little from the simulation, and those who seem to have gained much. The analysis suggests that simulation can provide a potentially powerful vehicle for experiential learning. The article concludes with a discussion of two sets of "working hypotheses," one associated with success in learning from simulations and the other with the various levels at which students function in experiential-simulation learning environments.

Pendant les années 1980, plusieurs programmes préparatoires universitaires pour les futur(e)s administrateurs et administratrices scolaires présentaient des pratica ou des cours genre-internats afin d'établir un meilleur équilibre entre la théorie et l'apprentissage pratique pour leurs étudiants et étudiantes. Dans cet article, on présente les résultats d'un projet qui examine le potentiel de l'apprentissage à but de formation administratrice en utilisant des présentations sur disques lasers interactifs sophistiqués. Il s'agirait d'une source alternative pour l'apprentissage expérientiel professionnel. On pose une question majeure: Dans un cycle d'apprentissage expérientiel, à quel point est-ce que la simulation peut-elle effectivement remplacer le véritable travail d'un directeur ou d'une directrice d'école? On cherche la réponse à cette question en examinant les expériences de deux sous-groupes distincts d'étudiants: un petit nombre qui semble avoir appris très peu de cette simulation et ceux qui semblent en avoir appris davantage. L'analyse des résultats suggère que la simulation pourrait être utilisée comme un moyen puissant pour l'apprentissage expérientiel. On

Bill Maynes is an assistant professor with the Department of Educational Administration. His current research interests include the study of experience-based learning in administrator preparation programs and the study of issues related to urban poverty and education.
Gordon McIntosh is a professor and graduate coordinator in the Department of Educational Administration. His current research interests include the study of leadership through the development of oral biographies of leaders in Alberta education and studies in the use of case and simulation methods in administrator preparation.
David Mappin is the Director of the Instructional Technology Centre and a doctoral student in the Department of Educational Administration. His current research interests include the design of human interaction simulations, multimedia design considerations, and organization and instruction in postsecondary education.

termine cet article avec une discussion de deux hypothèses possibles: l'une associée au succès de l'apprentissage par simulation et l'autre qui représente les différents niveaux de fonctionnement des étudiants dans des situations d'apprentissages expérientiels et simulatoires.

Recently, preparation programs for the school principalship have been the subject of renewed interest in universities across North America. In part this has been in response to the large number of influential studies that have made convincing claims that the principalship is the most important strategic role for school improvement (e.g., Fullan, 1982; Leithwood & Montgomery, 1982; Manasse, 1985). There is also evidence that teachers and school administrators see university studies at the graduate level as an essential qualification for the principalship (McIntosh & Montgomerie, 1987). Yet university-based preparation programs, despite their perceived importance and their widespread acceptance in the field, have been criticized as being "too theoretical." Manasse (1985), for example, recommended that "the structure of preservice training should allow students to experience first-hand the hectic, nonrational world of management through some kind of observation/internship/mentoring experiences" (p. 457). Similarly, the Report of the National Commission on Excellence in Educational Administration (1987) in the United States noted "troubling aspects throughout the field, including ... lack of preparation programs relevant to the job demands of school administrators, lack of sequence, modern content, and clinical experiences, in preparation programs" (pp. xvi-xvii).

In responding to this call for more experience-based learning opportunities, many universities have developed internship or mentorship-like programs (see, e.g., Ackerman, 1991 and Ratsoy, Richards, Haughey, & Maynes, 1992). Evaluations of these programs are generally positive and encouraging, but the same evaluations raise several cautions. For example, Ackerman (1991), based on his study of the Harvard Graduate School of Education Principal Certification Program, concluded:

> Inherent constraints place an unrealistic burden not only on the intern but on all parties involved in a preparation program. The limited time required for a practicum, variation in practices between sites, and inadequate supervision are only some of the factors which make the development of a full range of administrative "competencies" unrealistic. (p. 46)

In a previous article (McIntosh, Maynes, & Mappin, 1989) we argued that although internships and similar field-based programs in which students learn by doing are important components of preparation programs, the structure of opportunities in such programs presents a barrier to certain kinds of learning. Of particular concern in the *structure of opportunities* for principal interns is the inevitable trade-off between *breadth of experience* and *exercise of responsibility*. Breadth of experience refers to the range of issues that trainees will have the opportunity to work with and learn from. Exercise of responsibility refers to the extent to which trainees are given full responsibility for dealing with issues. This is the single most important factor contributing to the effectiveness of clinical training (Becker, Geer, Hughes, & Strauss, 1961; McIntosh, 1969). A different kind of learning results when trainees are respon-

sible for carrying out the work that they are learning to perform: they are pushed (or push themselves) to new levels of attainment and self-discovery.

Because of the potentially serious and farreaching consequences of actions taken in relation to many school-level administrative issues, educators who coordinate administrative field (or clinical) training programs tend to trade off responsibility against experience. Where trainees are given direct responsibility, they are given relatively unimportant administrative assignments; in the difficult situations actually faced by administrators, trainees tend to be involved only as observers, or they may simply be given reports as to what took place in critical encounters from which they have been excluded. They exercise no responsibility.

An obvious advantage of using simulation as a base for experiential learning is that students can be given complete responsibility to deal with a full range of administrative issues. There is no need for an "experience-responsibility trade-off." But simulation brings forth different learning issues, the most pressing of which relates to fidelity or the extent to which a simulation can become reality for those participating in the simulation. The recent development of interactive laser videodisc technology has been encouraging for developers working to address this issue. In this article, we report preliminary findings about student learning from one such simulation, a sophisticated interactive laser videodisc simulation of the principalship. We begin by presenting a discussion of the model, which from the outset has formed the basis of our teaching/learning strategies, then move to a brief description of the simulation. The remainder of the article focuses on how and what students have learned from their involvement in the simulation.

The Experiential Learning Model

Experiential learning is a conception of adult learning and professional development well described by Kolb (1984) in his book *Experiential Learning*, which has been developed over the past two decades. Kolb's conception is based on the idea that it is impossible and undesirable to segment the life of a career-oriented person into a section called *work*, a second section called *education*, and a third called *personal development* (see Figure 1). As Kolb said:

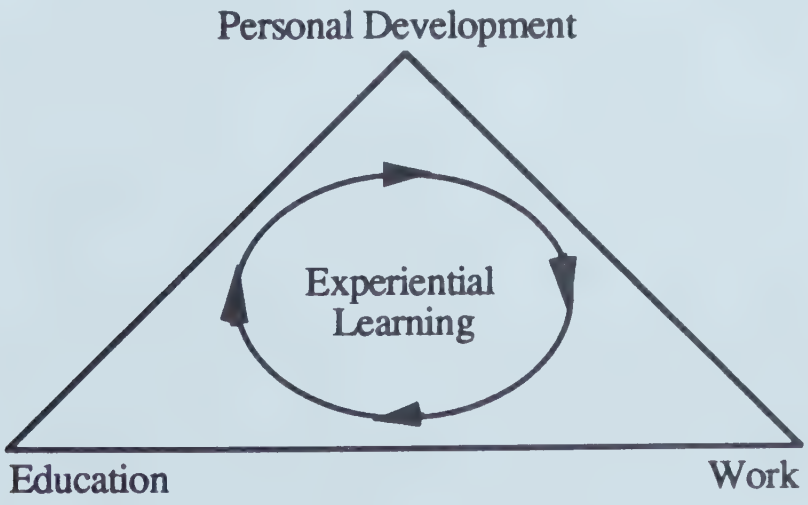


Figure 1. Kolb's conception of experiential learning.

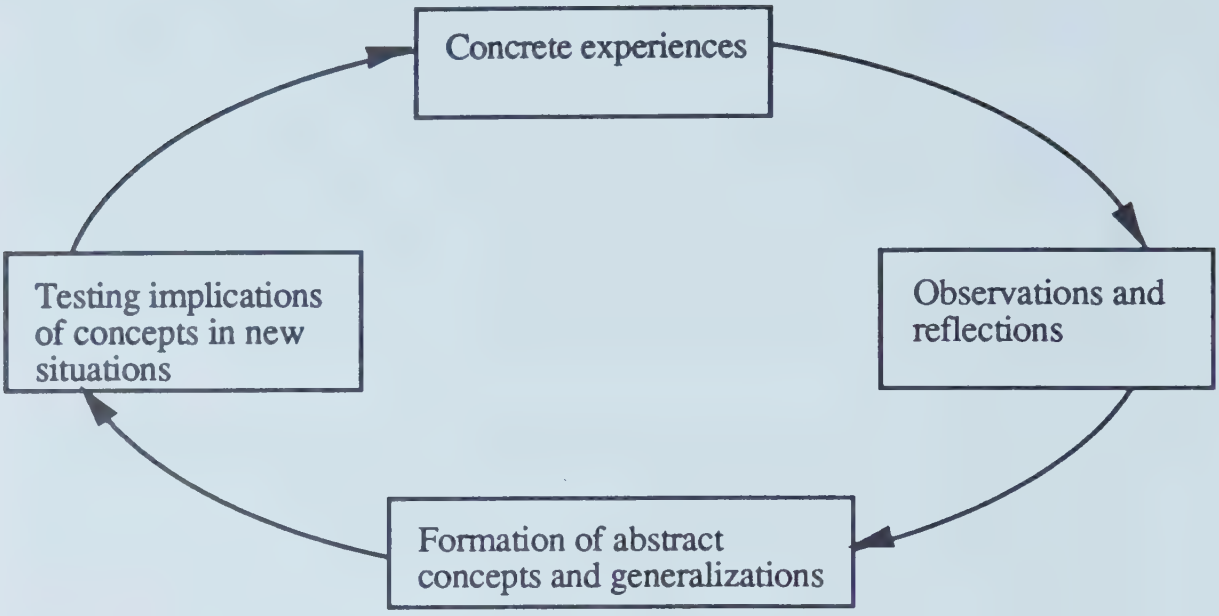


Figure 2. The Lewinian experiential learning model.

> It pictures the workplace as a learning environment that can enhance and supplement formal education and can foster personal development through meaningful work and career-development opportunities. And it stresses the role of formal education in lifelong learning and the development of individuals to their full potential as citizens, family members, and human beings. (p. 4)

The origins of experiential learning are derived from the work of John Dewey, Kurt Lewin, and Jean Piaget. Without presenting an extended explanation of their work and the ways their ideas are perceived to interact, it is possible for the present purposes to illustrate the basic idea of the experiential learning process by using a model derived from Lewin's work depicting learning as a four-stage cycle (see Figure 2). Kolb (1984) describes this cycle this way:

> Immediate concrete experience is the basis for observation and reflection. These observations are assimilated into a "theory" from which new implications for action can be deduced. These implications or hypotheses then serve as guides in acting to create new experiences. (p. 21)

The Experiential-Simulation Method

In bringing together the "disciplines" of experiential learning and simulation, we arrived at an instructional model with four essential elements:

1. experience of the field of practice, as provided by simulations of the work of the elementary school principal, in which the learner actively makes decisions—some of obvious major import, others apparently trivial—that are thrust upon the student-principal;
2. critical reflection on decisions taken and rationales offered for these decisions, first of all through discussion with peers and instructors in follow-up seminars, and later through reading and journal writing;
3. proactive planning in group tutorials based on assessments of the fundamental underlying issues characterizing the present situation in the

simulated school that must be addressed if the organization is to build on its strengths and best compensate for its weaknesses. These assessments are to be made within a framework of the student-principal's developing conception of the direction in which he or she, working with the staff, would like to take the school;

4. the organization of learning so acquired by the student-principal, integration of this personal learning with the literature pertinent to the field of study, and with personal values and capacities as these become progressively more clear to the learner.

A graduate-level course provides the structure in which this model is operationalized. From the first day of the course, students take on the role of Stacey Metzger, newly appointed principal of Pembina Elementary School. This aspect of the simulation is set late in June. The first task confronting students is to learn as much as they possibly can about the community, the district, and the school prior to school opening in September. The "magic" of interactive laser videodisc technology allows the presentation of a rich context within which the Stacey Metzgers will deal with administrative issues. For example, they talk with the superintendent and the outgoing principal, they have access to a relatively complete set of school files, and they see all of their teachers in the classroom interacting with children, then in other areas interacting with colleagues. That all participants in the simulation will deal with the same issues within the same rich context represents a significant advance on previous approaches to simulation or case study. But the three simulated half-days of work provide the focus for the most significant of the learning activities.

By not later than the fourth class session, the student-principals face their first half day in the role of principal. This first experience on the "firing line" takes place, according to the scenario stipulated for the exercise, in early September as the school year is getting under way. (The two subsequent exercises are set in early November, at the time of the first reporting period, when the principal could reasonably be expected to have gained a good sense of the state of the school; and in early spring, at the time when the budgeting/planning process for the next school year is well under way.)

As they begin their work day, the school secretary appears on the screen, welcomes the principal back to the school after the two-day absence stipulated in the scenario for the exercise, and provides a briefing on several concerns that have come to her attention. This is followed by a similar briefing from the assistant principal. Then a menu appears on the screen that lists the various items that have accumulated in the principal's in-basket. The student principal has 135 minutes to address the set of problems, which number in the range of 23 to 25 for each of the three half-day work sessions.

Problems are presented to the student-principals in several ways: by video presentation of paper items such as notes, memos, and other forms of correspondence; by video presentation of visitors to the principal's office (who are invited in by the principal); by unscheduled interruptions (not controlled by the principal); and by telephone calls.

The student-principals use an electronic reporting form to record their decisions and actions during the work session. At the end of the half-day's

work, the student-principals print out the record of their decisions, which they bring to the follow-up seminars and group tutorials.

The seminar, group tutorial, reading, and writing aspects of the course are important complements of the laboratory component. In turn, these latter activities, engaged in outside the simulation laboratory, are dependent for their significance on the experience students gain of the work life of the school principal in the simulation laboratory. The various course activities, in short, are mutually interdependent.

Based in major part on earlier studies of experiential learning, and to a lesser extent on studies of case and simulation methods, we anticipated that through the experiential-simulation method, students would learn about: (a) the personal demands of the principalship, especially those demands that are not meaningfully communicated "by the book"; (b) one's personal "fit" with the demands of the role, that is, the degree to which the work of the principalship can be brought into a comfortable alignment with the personal strengths, values, and aspirations of the student-principal; (c) management and human relations skills, in particular, conflict management and time management skills; (d) the development of personal resilience and sensitivity to monitor one's role performance and identify directions for improvement; and (e) problem analysis capabilities.

Studies of case and simulation methods offered some support for the expectation that problem analysis skills could be developed using experiential-simulation methods (e.g., Fisher, 1978). But the likelihood that the remaining expectations could be achieved rested on an untested assumption, namely, that the "work" component in the Kolb experiential learning model could be successfully replaced by simulations of work experience. We set out to test this assumption by addressing the question: What do students learn when they are engaged in the experiential-simulation mode of instruction?

Student Learning

Over the past 10 years of continuous work on this project we have addressed three kinds of research and development issues: technical issues (e.g., how can we best use the available microcomputer and videodisc technology to make the learning environment more interactive or responsive to the actions of the learners?); pedagogical issues (e.g., what are the inherent limitations of substituting simulations for direct experience of the world? what are the inherent limitations and strengths of "permissive" learning procedures?); and evaluative issues (e.g., what are students learning in this course? what serves to facilitate and impede student learning?).

In practice we have found that the issues are interwoven and that the fundamental issue is expressed in the evaluative question: What are students learning? Over the past five years, we have been examining this question through analysis of student journals and observations of laboratory and classroom behavior. In the analysis that follows, we discuss the approaches to learning, and the substance of the learning, of two distinct subgroups: those who seem to have gained little from their involvement in the simulation, and those who have gained much. Although the subgroup of those who gained little is quite small, their experiences of the course are quite informative. Analysis of data collected from this group has led us to propose a fourfold

typology of resistance to learning in the experiential-simulation mode. Here we present this typology in summary form so that it can serve as a contrast to our more extensive analysis of "model learners." A more thorough discussion of this typology can be found in McIntosh (1989).

Resistance to Learning in the Simulation-Experiential Mode

Two criteria guided the selection of the students on whose experiences this analysis was based. The first was that in what was written in their journals there was little evidence of personal growth. The second was that our subjective assessments based on observations led us to conclude that they had learned little. Our analysis suggests four identifiable types of resistance to learning:

The *intellectualizer* is unable or unwilling to relate the experience of the simulation laboratories to reading in the research literature on the principalship, and to relate both of these to his or her personal development as an administrator. The intellectualizer is unable either to understand or to accept the transformational role to be played by the experiential learner—the learner as creator of personal knowledge.

The *active resister* has a clear, even shrewd, perception of the problems facing the school principal but is unable or unwilling to consider courses of action that differ from the tried and true methods arrived at on the basis of previous experience. He or she is dismissive of alternative courses of action suggested by peers or course instructors in seminars and group tutorials. The active resister says, in effect, "I cannot change. I am no longer open to the consideration of alternatives."

The *distorted interpreter*, like the active resister, is in thrall to past experience, but the nature of the resistance is different. This learner's perceptions of reality seem to be distorted by events experienced in the past. For example, one of our students came to us from a recent unsuccessful experience as a school principal. His "distorted vision" was most evident in any situation where staff involvement in decision making was discussed. In different ways he would express his views on staff involvement in terms such as these: "Administration is a struggle for power.... You've got to let them know who's boss.... Teachers are always out to get the principal." Thus, personal history acts as a filter that often interferes with the distorted interpreter's perception of problems that he or she is called upon to address in the laboratory setting.

The *self doubter* may be keenly perceptive and avidly open to the style of learning called for in the course. However, this learner brings to the experiential learning setting a legacy of self doubt. The following pattern characterizes one self doubter we have worked with in the course: rigorous, almost painful, honesty in self assessment of performance in the simulation exercises; a propensity to "get down" on himself; a tendency to dismiss his own analyses as insufficient authority for resolute action; an overwillingness to second guess his actions as he reflected on his performance; and a penchant for engaging in agonizing reappraisal of beliefs and commitments. We see the self doubter as someone open to further learning, but because of self doubt unable to take full advantage of the insights gained from experience and thus unable to develop as a person on the job.

Model Learners

In the discussion of resistance to learning, our purpose is to examine the experiences of our least successful course participants so as to inform our view of how students learn in simulation environments. In this section, we pursue a similar purpose, but through an examination of the experiences of two successful learners.

Elaine (a pseudonym) has had 10 outstanding years as a high school teacher, and at the time of her studies with us was on sabbatical leave to complete a master's degree in educational administration. She has had no administrative experience and at the beginning of the course was not certain whether she wanted to become a school-level administrator. Her immediate career goal was to become a consultant with the central office in her district. One of Elaine's major purposes in taking the course was to arrive at a personal resolution with respect to whether she would pursue an appointment as a school-level administrator.

Steve (a pseudonym) is an experienced and successful school superintendent who was also on sabbatical leave at the time of his studies with us. He holds a Master of Education degree, and was registered in a postgraduate diploma program in the Department of Elementary Education. Steve felt that because all of his teaching and school-level administrative experience was at the secondary level he was not as in touch with the elementary schools in his district as he would have liked. He therefore set the goal of his sabbatical leave as learning as much about elementary education as possible, hence his choice to register in our course which simulates an elementary school principalship.

Elaine

Elaine approached the course with a tremendous amount of self-doubt, which she expressed in her first journal entry as follows:

> This course both excites and frightens me.... Much of this fear comes from my perceived inability to handle day-to-day decision making that affects so many, as is the case in the principalship. The fear of making the wrong decision and losing face with staff, parents, and students is very real for me.

Then later in the same journal entry:

> As I ponder, the more I question my ability to be a principal. I am excellent at organization, but have difficulty asking others for help. I can see areas in need of change but would have difficulty persuading others because of my lack of self-confidence. I am a sincere person and would have difficulty reprimanding others. I am a perfectionist and worry about burning out.

Her self-doubt is obvious, but in the latter quote, we see another dimension of Elaine. She takes careful measure of her strengths and weaknesses and is open to admitting both. Throughout Elaine's journal we see evidence of a strong internal locus of control. She takes full responsibilty for what she is and for what she can become. This factor in itself may explain why Elaine's self-doubt proved not to be disabling as it was in the case of the self doubter we presented earlier in this article.

As Elaine worked through her self-doubt, she found that much of it was rooted in an elevated view of what good principals do and are capable of doing. Having a superintendent and several experienced principals present in

seminars to discuss how they had chosen to deal with specific incidents proved to be liberating for Elaine. To her it was a revelation that principals do not always have a "one best answer" to a situation. About this she wrote:

> It is very reassuring that persons currently in principalship roles don't necessarily agree on strategies they would use in each case. Also, many have specified that their decisions would change depending on the context of the situation for each child. In other words, there are no right or wrong definite decisions. As a math person, this is hard for me. I am looking for a formula whereby if this happens you do this—every time! ... I am now beginning to see that people just don't fit into formulas and so can't all be treated exactly the same. *I'm also getting the feeling that a principal is only human, bound to make mistakes, and must get up, shake it off, reaffirm their belief in themselves, and carry on.* [our emphasis]

She concluded this reflection with a powerful quotation from Eleanor Roosevelt: "No one can make you feel inferior without your consent."

As the course unfolded, Elaine included more and more comments about her ability to do a good job as a principal. She wrote that the exceptional principals she knew worked at the "symbolic" level and:

> Their very actions spoke their values and beliefs of schooling. *I think I would be able to do this.* I always try and "practice what I preach" to my students, and it gains their respect. [our emphasis]

When Elaine discovered that in many cases she made the same decisions as practicing principals, her confidence was further enhanced. After each simulation experience, students have access to descriptions of how the principals in real life handled the situations on which our simulation is based. Following the second simulation experience, Elaine wrote:

> I checked the actual principal solutions for the items in In-basket #2. I find this activity completes the cycle for me, and I find it especially useful after the debriefing session where the class has had a chance to share possible solutions. *I was pleasantly surprised that my actions coincided with many of the "actuals."* [our emphasis]

Even as she was gaining confidence, Elaine continued to be self-critical, but never in a disabling way. Her self-criticism was aimed at learning. This is evident in what she wrote about one of the incidents presented during the simulation in which a male driving a blue Camaro had attempted to pick up two girls on their way to school:

> The Blue Camaro incident was a real shock for me as we debriefed it. All I had considered doing was following up on the police investigation within two days. I assumed that because the police were involved, I had no other responsibility. What a mistake! What a learning experience! That is one reason why I feel a course like this is so important. I don't want to learn about being a principal the hard way—at least, not out in the real world. I think it is valuable in this kind of setting where you can make mistakes the hard way and know it is only intended to teach. I have a real fear of losing credibility as an administrator, and an incident like this could easily do it. Now I know how I might handle it differently and more effectively.

There are many examples in Elaine's journal of her openness to this kind of learning. She continuously reflected on her performance in the simulation.

And most often it would have been difficult for an uninformed reader to know that she was not writing about real experiences. The extent to which Elaine projected herself into the role of Stacey Metzger is reflected in a comment she wrote in her journal the week prior to a simulation experience. She had just made a very confident statement about her ability to "handle this kind of position." Then wrote: "So why is my stomach still churning in anticipation of next week??" We believe that it was churning because Elaine saw herself as the principal, responsible both for resolving the issues with which she was to be confronted next week, and for setting a direction for her school. There is no evidence that she made any effort to distance herself from the simulation.

The way Elaine worked toward developing a personal understanding of the events of the course caught our attention. In almost all of her journal entries she brought to bear her own views, the views of others in the course, and views expressed in the literature. Her resolution of whatever issues she was grappling with inevitably was an amalgam of these perspectives. This suggests to us that she viewed all three sources as valuable and important to her search for understanding. Compare this with the self doubter who dismissed his own views, and the active resister who dismissed any view other than her own.

Elaine's work ethic also deserves mention. Her journal contains reference to most of the readings that students were invited, not required, to read. She spent hours learning about the context of the simulation—the school district, school, students, and staff of Pembina Elementary School. And her journal is a small book!

What was Elaine's final resolution with respect to whether she ought to pursue a school-level administrative position?

> [The course] has helped me tremendously with my self-confidence in a way that theory classes aren't capable of. Role playing is a definite advantage of this course and I would recommend it highly to anyone having no administrative experience. It has further helped me decide that I *can* handle such a position and that I am willing to try it. [Her emphasis!]

Steve

The first task faced by students in our simulation course is to become familiar with the context in which they will be working. And this is no mean feat. There is a very large quantity of information to be mastered about the school, the district, the students, and staff. Steve approached this task in the way we intend; he assumed complete personal responsibility for knowing this information. And, as would a new principal moving into an unknown school district, he extended great effort to acquire that knowledge. This is just one example of the extent to which Steve projected himself into the role of Stacey Metzger, principal of Pembina Elementary School. Steve's learnings from the effort he extended at this stage of the simulation are interesting. He wrote:

> I've just completed a four hour session of going through the Filebox/Orientations/ ... to better familiarize myself with Rutherford and Pembina and am suffering the effects of "information overload." On a number of occasions during my career I've experienced the stress and overwhelming need of learning a vast amount of information about a school's or organization's circumstances in order to be able to carry out the requirements of a new job, but I guess that

> time has dulled my memory. This simulation has reactivated the feelings (memory) in vivid clarity.

This experience reminded Steve of the complexity and difficulty of the principal's job:

> A point reemphasized to me by the simulation was that principals have to know a hell of lot about a hell of a lot of things. Their jobs are very demanding and cover a large gamut of skills and range of knowledge about many areas.

These were significant insights with consequences for action. Steve has now planned a much more accessible and supportive program of orientation for neophyte administrators in his district. Of his learnings, he wrote:

> I know that the foregoing does not represent any real profound insight, but it has really impacted on me—maybe in a "feeling" (emotive) rather than "thinking" (cognitive) way. Whatever, it has motivated me to resolve to be more proactively supportive during the beginning period of a newly appointed administrators tenure.

In response to this comment, we wrote: "Profound insights are personal phenomena, not to be measured by any standard outside of your own experience." Steve's insights illustrate the way he personalized the simulation experiences and utilized them to examine current practices and beliefs, leading to personal insights and commitments to action.

At the beginning of the course, we ask students to write to us in their journals about what they hope to gain through their participation. This is part of Steve's response:

> I'm particularly looking forward to being able to respond to simulated issues (In-Basket exercises) and then analyze and compare my responses to those of classmates. We are unable to do many meaningful comparisons in real life—there, we're the only ones who took the action and to ask another person what they *would have* done in our place often gets just an offhand opinion (the issue is hypothetical to them and consequently they will not put the soul searching into it that you did).

Indeed, as the course unfolded Steve did find this a rich source for reflection. The range of experience represented by the students in the class contributed to this. There were teachers, consultants, assistant principals, and principals with whom Steve (the superintendent) could examine decisions taken during the simulation. He commented a number of times that it was useful for a superintendent to revisit teachers' and principals' perceptions and approaches to specific incidents and to broad educational issues. Indeed, this was the source of some personal reflection. With respect to one set of staff-related simulation incidents, he commented:

> My suggested courses of action ... when compared to those of other class members reveal that I was inclined to adopt a more direct and abrupt response.... It appears from the comments of my classmates that I was inclined to "shoot faster" than they were in these instances, and this prompts me to consider the pros and cons of my stance.

In the consideration that followed, Steve examined the potential strengths and weaknesses of his approach. Like Elaine, he openly wrote of both. There is

no evidence of defensiveness, only of openness to learning. For Steve and Elaine, self-criticism appeared to be a catalyst to learning. This is in sharp contrast with the disabling effect that self-criticism had on the self doubter.

Steve's openness to learning is well illustrated through his experiences with the telephone-caller component of the simulation. The telephone evaluation reports that Steve received after the first simulation experience suggested a level of performance that did not meet his standards. Based in part on our seminar debriefing of the telephone call incidents, Steve set in place a plan to improve his telephone performance during the second simulation experience. He reported that as a consequence of his plan, "my overall rating, and those of most categories have improved." But he was concerned that "I still have a problem of working to facilitate an understanding of the nature of the caller's problem and its importance to him/her." Reflecting on this, he wrote:

> I realize that two sets of three phone calls do not provide definitive proof regarding the nature of my telephone skills but they may very well suggest valid indications (what the heck, they represent the best (and only) data I've got on the issue). I'm heartened by the indicated improvement, and will continue to work on this skill.

In this endeavor, as was the case with all of the matters with which Steve dealt during the course, his search was not limited to an exploration of personal experience. He read voraciously and incorporated what he read in his thinking. In a journal entry written four days after that quoted immediately above, Steve reported having read an article entitled "What to Do When You're Verbally Attacked." Of this he wrote:

> This article has suggested one possible source of my problem. I think I try to impede the caller's anger/criticism rather than trying to understand it; I may, too quickly, try to maneuver the conversation away from statement of the problem (negative message) in an effort to get to defence (blunt the attack) or solution. By my doing so, the caller probably feels that he hasn't "had his day in court" and that I'm not really interested in his concern or its importance to him. I'll have to try to slow down and give the initial outpouring its due time.

This was Steve's typical approach to thinking about course experiences. He moved easily from using the simulation experiences as a filter for his thinking, to using his superintendency experience as a filter, to using the literature as a filter. And he was comfortable in any combination of these approaches. Steve is a sophisticated learner.

Steve, like Elaine, demonstrated an exceptional work ethic. One of the reasons that he was able to integrate the literature so well with his thinking is that he read all of our suggested readings and sought out other sources on his own. And his journal is a larger book than Elaine's!

What has been written above does not adequately describe Steve's sophistication as a learner, but perhaps it does convey enough to provide readers with a sense of why he was successful in the course. The last words about his success in the course go to Steve:

> Aside from the insights I've gained from reading, reflection and class interaction; and the improvement of skill (hopefully my telephone skills); I've gained a reintroduction to and appreciation of, the principal's roles, pressures and

> frustrations. I think that this will have a positive impact on my interactions/decisions/relations with principals; both they and I, and consequently the system, will benefit from this.

Reflections on Success

The approaches to the course of Elaine and Steve (and others who were successful in our course) had much in common. These commonalties, juxtaposed with observed characteristics of those who have demonstrated resistance to learning, have led us to seven propositions about factors likely to contribute to success at learning from simulations. We think of these as "working hypotheses" that we will continue to explore.

Work Ethic

Both Elaine and Steve demonstrated an exceptional work ethic. They put in hours beyond what would be required for satisfactory performance in the course, and this without the external motivator related to the possibility of obtaining high marks (the course is graded credit/non-credit). We speculate that they bring this same work ethic to all that they do.

Ability

Elaine and Steve are talented, intelligent, and thoughtful. Perhaps Griffiths (1988) was correct when he wrote, "there are no recorded examples of good dumb principals or successful stupid superintendents" (p. 12).

A Base of Successful Practice

Elaine was an accomplished teacher, and Steve a successful superintendent. The self doubter, on the other hand, had experienced considerable difficulty in his latest placement as a teacher. In part this may explain why Elaine was able to use her self-doubt as a catalyst for learning, while the self doubter was disabled by his. The student with distorted vision provides a second example of a person whose unsuccessful practice interfered with his ability to learn from the course experiences.

Internal Locus of Control

Elaine and Steve demonstrated strong internal loci of control. They accepted responsibility for what they were and for what they were to become. This left them searching for, and open to, learning opportunities.

Purpose

Although Elaine's and Steve's purposes for taking the course were different in some respects, they were similar in fundamental ways. Both sought to learn about the principalship from a perspective that was very personal and significant with respect to the positions they held in their school districts. Elaine wanted to know whether she could handle a principalship, and Steve wanted to better understand the work of the elementary school principals he supervised. Their approach was in marked contrast to that of another student who appeared to be taking the course only to fulfill the requirements of a master's degree. This student seemed to always know the minimum amount he needed to do for credit. In pursuing that minimum, he found it! He gained very little from the course, but then, that was his purpose.

Valuing All Sources of Information

Elaine and Steve sought and carefully considered information from themselves, colleagues, instructors, and the literature. The self doubter, on the other hand, did not value his own views, while the active resister valued only her own views. Neither approach facilitates learning through reflection.

Becoming Stacey Metzger

Elaine and Steve projected themselves completely into the principalship of Pembina Elementary School. They assumed responsibility for what was happening in the school, and became affectively and cognitively engaged in their roles. The decisions they were making were real to them, and they were responsible for the consequences of those decisions. This is what the intellectualizer was unable to do, and what resulted in her distant and academic approach to the course.

Levels of Experiential-Simulation Learning

Based on early work with the simulation course, McIntosh (1989) proposed the following as a framework helpful in understanding the levels at which students worked in our experiential-simulation learning environment:

Level 1. The learner understands the learning model and accepts responsibility for transforming experience into personal knowledge as called for in the model.

Level 2. The student is able to discriminate the significant from the insignificant in the experience which he or she encounters. This might be referred to as the "ah-hah" stage. The student has capacities for sensitive perception, for clear-minded recognition that a particular experience is an opening toward learning more about his or her performance of the principal's role.

Level 3. The student is able to work with an intuition that an experience is important, and bring to conscious awareness by valid interpretation the significance of the experience (or passage from the reading).

Level 4. The student establishes connections between seemingly disparate and unrelated experiences so as to discern "unifying strands of insight." The knowledge so acquired is personal knowledge, that is, it is knowledge of self in relation to knowledge of the roles, relationships, and dynamics of organizational behavior in schools and the administrative tasks of schooling.

Level 5. The fifth level is that of personal resolve and action, that is, of commitment to act upon the new learning in relevant ways and to reflect upon the effects of these actions.

We do not view this model as describing a sequence through which all learners move, nor do we yet understand the extent to which the model will be useful in helping us understand the development of individual learners. It is useful, however as a heuristic for understanding the levels at which a particular learner is operating at a particular time. And we have observed that the most sophisticated of our learners seem to function regularly at the higher levels, while less sophisticated learners very seldom work at levels 4 and 5. At this early stage in our work we view this model of experiential-simulation learning as tentative, as a set of working hypotheses about which much more is to be learned. The analysis of model learners, however, increases our confidence in the model. The descriptions provided above demonstrate that

Elaine and Steve did engage in all five levels as they learned from their simulation experiences. Moreover, *they* seemed to define their success in terms of whether they were able to arrive at some sort of personal resolve and commitment to action (level 5).

Our understanding of the model has been enriched by the analyses, both of resistance to learning, and of model learners. The former allows us to speculate on possible roadblocks that may confront students in their efforts to learn from our course, and the latter to speculate on characteristics students possess, or may choose to invoke, and that may enable them to successfully negotiate each of the levels of learning.

We also feel closer to an an understanding of one of the central issues related to learning from simulations. For us, this issue is expressed in the question of whether in our course simulation effectively replaces work in Kolb's conception of experiential learning. We no longer expect an absolute answer to the question. Certainly simulation cannot become real work. But in relation to our instructional purposes, and for some students, the simulation acts pedagogically as the functional equivalent of work. Students like Elaine and Steve project themselves fully into the role of Stacey Metzger. For them, the simulation is work. Other students have distanced themselves to some extent from the role. For them the simulation is, in various degrees, less like work, less like the real thing. We now look at this aspect of experiential-simulation learning as a type of contract. Our side of the contract is to provide as inviting a simulation of the work life of a principal as possible. But the contract is engaged only at the will of the student. We believe that the analysis reported in this article brings us closer to understanding the nature of the contract and our role in facilitating engagement.

References

Ackerman, R.H. (1991). *Portraits of practicum life: Toward an understanding of field experience for aspiring principals.* Paper presented at the annual meeting of The American Educational Research Association, Chicago.

Becker, H.S., Geer, B., Hughes, E.C., & Strauss, A. (1961). *Boys in white: Student culture in medical school.* Chicago: University of Chicago Press.

Fisher, C.F. (1978). Being there vicariously by case studies. In O. Milton (Ed.), *On college teaching* (pp. 258-285). San Francisco: Jossey-Bass.

Fullan, M. (1982). *The meaning of educational change.* Toronto: OISE Press.

Griffiths, D. (1988). *Educational administration: Reform pdq or rip.* UCEA Occasional Paper, Tempe, AZ.

Kolb, D.A. (1984). *Experiential learning: Experience as the source of learning and development.* Englewood Cliffs, NJ: Prentice-Hall.

Leithwood, K.A., & Montgomery, D.J. (1982). The role of the elementary school principal in program improvement. *Review of Educational Research, 52,* 303-309.

Manasse, A.L. (1985). Improving conditions for principal effectiveness: Policy implications of research. *The Elementary School Journal, 85,* 439-463.

McIntosh, G. (1969). *A comparative study of clinical training.* Unpublished doctoral dissertation, Harvard University.

McIntosh, G. (1989, March). *Learning the practice of school administration.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

McIntosh, R.G., & Montgomerie, T.C. (Eds.). (1987). *The status of the school principal in northeastern Alberta.* Edmonton: University of Alberta, Department of Educational Administration.

McIntosh, R.G., Maynes, W.G., & Mappin, D.A. (1989). Preparation for professional practice: A case approach. *The Canadian Administrator,* 28(7), 1-5.

National Commission on Excellence in Educational Administration. (1987). *Leaders for America's schools.* Tempe, AZ: University Council for Educational Administration.

Ratsoy, E., Richards, D., Haughey, M., & Maynes, W. (1992). Becoming a principal, part 1: An innovative preparation program. *Canadian School Executive, 11*(9), 3-6.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 4, December 1992, 285-299*

Paul T. Begley
Ontario Institute for Studies in Education

and

Glenda Campbell-Evans
Edith Cowan University

Socializing Experiences of Aspiring Principals

The research addressed the socializing experiences of aspiring school principals enrolled in a mandated principal certification program. This was done in an effort to improve our understanding of certain socializing influences, not necessarily related to program content, that affect the impact of formal principal preparation courses. Data were collected by means of a pre-course survey, interviews, and dialogue journals from a sample of 87 candidates enrolled in a preparation course delivered in the Northwest Territories. An original conceptual framework, derived from earlier work by Van Gennep (1960) and Greenfield (1985), as well as a grounded analysis procedure, were employed as data analysis procedures. Results suggest that internal processes (e.g., personal challenge, thirst for knowledge) figure prominently as influences on individuals choosing to pursue school leadership roles. Organizational/contextual factors (e.g., a mandated course, financial support), and the influence of other people, become influences somewhat later in the socialization process.

L'objectif de cette recherche s'adresse aux expériences de socialisation de futurs administrateurs et administratrices scolaires inscrit(e)s au programme de certification obligatoire. Le but de cette recherche vise à améliorer notre compréhension de certaines influences sociales qui n'étaient pas forcément reliées au contenu du programme mais qui influencent l'effet voulu de ces cours d'administration préparatoires formels. Les données ont été collectionnées par l'entremise d'un questionnaire administré avant la présentation du programme, par des entrevues et par un journal où on dialoguait avec des participants et participantes choisi(e)s de 87 candidat(e)s inscrit(e)s au cours préparatoire donné aux Territoires du Nord Ouest. L'encadrement conceptuel a été inspiré d'un travail réalisé au préalable par Van Gennep (1960) et Greenfield (1985) ainsi que par des procédures d'analyses de données actuelles qui servaient de moyens d'analyse de toutes les données. Les résultats suggèrent que des processus internes (par exemples : défis personnels, soif de connaissance) sont des influences très importantes dans la vie des individus qui choisissent de poursuivre des rôles administratifs scolaires et des rôles de leadership. Des facteurs contextuels et organisationnels (par exemples: cours obligatoire, appui financier) ainsi que l'influence subit d'autrui, deviennent aussi des influences qui apparaissent, bien qu'un peu plus tard.

Paul T. Begley is an associate professor at the Midwestern Field Centre. His areas of academic interest in the Department of Educational Administration include administrative values, school improvement processes, and leadership development in cross-cultural environments.
Glenda Campbell-Evans is a lecturer at the Department of Educational Policy and Administrative Studies. Her research efforts are currently focused in the area of the professional development of practicing and aspiring school administrators.

Departments of education in North America and Australia continue to rely heavily on formal training programs as the chief means of socializing aspiring school administrators. There also appears to have been a general shift away from the more serendipitous and informal socialization processes typical of the past (Feldman, 1989; Schein, 1968; Van Maanen, 1976), and an increased commitment toward specific and often mandated preparation experiences.[1] This is a trend that has been supported, and indeed fueled, by the substantial body of applied research that has accumulated in recent years purporting to describe the nature of administrative expertise (see Burdin, 1989; Greenfield, 1987; Leithwood, Begley, & Cousins, 1992).

Principal preparation programs can now be grounded in more than just the context-bound practices, or theories-in-use (Argyris, 1982) manifested by local practitioners serving as instructors. The availability of research-derived knowledge on effective administrative practices has permitted the design and development of much better validated preservice programs for principals (see Daresh & Playko, 1992). However, as positive an outcome as this might seem, we still don't know much about the factors that cause individuals to pursue school leadership positions and enroll in principal preparation programs in the first place. Furthermore, as Leithwood, Steinbach, and Begley (1992) report, the disturbing reality is that "formal preparation programs for aspiring and practicing administrators appear to vary widely in their perceived value ... extremely helpful or not very helpful at all" (p. 301). Apparently a number of factors not necessarily related to program content influence the impact of formal principal preparation courses.

Trainers and sponsors of school leadership programs have discovered the need to be sensitive to the influence of certain other factors associated with the socialization of both aspiring and incumbent administrators. For example, in our experience as program developers since 1985, we have noted that people enroll in our training programs for a variety of reasons and with a range of expectations. Readiness levels can vary so greatly within a given cohort of candidates that course implementors are sometimes presented with instructional challenges that parallel, or even exceed, those faced by a teacher in a completely destreamed elementary panel classroom. To be specific, depending on whether a serious needs assessment process precedes course development and what criteria are employed for candidate selection, participants enrolled in principal preparation programs may bring to the experience a relatively broad spectrum of prior experiences, training needs, and personal expectations. These "base rate" characteristics (Feldman, 1989) constitute candidate readiness for participation in a particular program. A typical outcome is that the needs of all candidates may not be met by the provision of experiences that instructors have identified as conceptually relevant, no matter how exemplary the syllabus. Given our apparent reliance on formal training programs as the primary means of socializing aspirants to school leadership roles, this is cause for some alarm. These issues of divergent expectations and varying levels of candidate readiness could be simply accepted as confounding variables that must be controlled for by instructors during course delivery. However, we preferred to view the situation as an invitation to conduct further research on the nature of those factors that influence the socialization

of aspiring school administrators. In particular we were interested in discovering how individuals decide to pursue school leadership roles and what motivates them to enroll as candidates in preservice training programs. This article reports the findings of our initial efforts in this direction.

Objectives of the Study

The context of our inquiry was a principal certification program developed by the Centre for Leadership Development at the Ontario Institute for Studies in Education under contract with the Northwest Territories (NWT) Department of Education. The general intent was to conduct research on the factors influencing the socialization experiences of approximately 100 aspiring and incumbent principals prior to and during their participation in a two year preservice program. Three specific research questions were addressed by the study:

1. For what reasons do course candidates decide to pursue the principalship as an educational role?
2. What influences do course candidates report on their decision to enroll in a formal preservice program intended to prepare them for a school principal's role?
3. What factors influenced the candidates' decisions to enroll in the Northwest Territories Principal Certification Program?

Alternate Images of the Principalship

Research findings on school leadership practices are now quite well documented (e.g., Leithwood, Begley, and Cousins, 1990). A number of well validated conceptual frameworks have emerged. Among the most notable are those that focus on the notions of instructional leadership, school improvement and, more recently, transformational leadership (e.g., Duke, 1987; Greenfield, 1987; Hallinger & Murphy, 1987; Leithwood & Steinbach, 1991). More to the point, we now have an increasingly firm basis on which to describe the knowledge, skills, and attitudes important to school administrators. This is not an insignificant achievement given the lack of consensus typically encountered on the questions of purpose and goals for school administrator preparation processes (Begley & Cousins, 1990).

In Ontario, the Northwest Territories, and Western Australia virtually all principal preservice preparation programs now manifest varying degrees of commitment to the view of the principal's role usually described as instructional leadership. It is possible to argue convincingly in support of more recently developed conceptions of the role, such as transformational leadership. However, given the mandated nature of the syllabus for most principal certification programs, and the natural time lag encountered when new knowledge impacts on established procedures, we concluded that instructional leadership remained the appropriate image of the role on which to base our inquiry.

Instructional leadership, the desired goal of socialization for our immediate purposes, is usually characterized by the presence of a clearly articulated educational philosophy, extensive knowledge about effective educational practices, and a clear understanding of the policy environment framing the school's purposes and practices. This image of the role emphasizes leadership

as well as management functions and extends considerably beyond the performance of routine work tasks, the usual domain of training research. In the research community, justification for an instructional leadership approach to school administration has been based largely on the premise that certain characteristic actions of principals, intended to encourage and support classroom practices linked by research to improved student outcomes, have a positive impact on student achievement. However, research by Heck, Larsen, and Marcoulides (1990) has also begun to validate the instructional leadership role of principals as a causal link to improved student outcomes. The characteristic actions associated with instructional leadership that have been positively linked to student achievement include managing the political relationship of the school to its environment, supervising the school's instructional organization, and building a positive climate for learning.

To foster such an orientation among aspiring principals, particular training metaphors, typically derived from research, are employed to structure preparation programs. For example, the principal preparation programs sponsored by the Ontario Institute for Studies in Education (OISE), delivered annually in Ontario and the Northwest Territories, employ Leithwood's Principal Profile (1988) as a basis for describing and articulating an instructional leadership role along four dimensions (goals, factors, strategies, and decision making) and four levels of role performance ranging from typical practice to exemplary practice.[2]

A Role-Specific Conception of Principal Socialization

Feldman (1989) sheds some conceptual light on the subject of socialization by reviewing several definitions in a chapter-length piece that compares the contribution of socialization research and training research to organizational development efforts. He cites Brim's (1966) definition, which we find particularly apropos, "the manner in which an individual learns that behavior appropriate to his position in a group through interaction with others who hold normative beliefs about what his role should be and who reward or punish him for correct or incorrect actions" (p. 3). Feldman (1989) also points out that socialization research has tended to focus most closely on the informal ways by which individuals learn about their work settings, and the impact of others on that learning. Training research, in contrast, has been most concerned with the ability of individuals to perform routine work tasks (p. 377).

In response to Feldman's observations, we decided to integrate both the socialization and training perspectives in the conceptual framework of our study. To achieve this integration, we chose to assess the influence of various socializing experiences on aspiring principals in terms of the contribution they make toward the attainment of a particular espoused image of the effective principal. For our purposes that image was the instructional leader.

Basic to our inquiry was the need for a role specific conception of the factors that at least hypothetically leads to the development of an instructional leadership orientation to the principal's role. We were also interested in recognizing, within our framework, the developmental aspects of the socialization process. Although the literature in this area of socialization research is not extensive, work by Van Gennep (1960), Greenfield (1985), and Leithwood (1988) were found to be most useful.

Factors	Stages of Socialization (Van Gennep, 1960)		
	Separation	Transition	Incorporation
Internal Processes:			
- values			
- cognitive processes			
Relationships with:			
- superordinates			
- peers			
- subordinates			
Organizational/Contextual:			
- organizational culture			
- district policies			
- selection/promotion procedures			
- formal training programs			
- informal training (mentoring etc.)			
- communication networks			
- planned critical events			
Image of the Role:			
- goals			
- factors (school/classroom)			
- strategies			
- decision making			

Figure 1. Conceptual framework (Begley 1990)
Factors contributing to the development of principals with instructional leadership orientations.

Van Gennep's (1960) three-stage model of professional socialization (separation, transition, and incorporation) proved helpful in conceptualizing the developmental process whereby individuals are influenced by a number of socializing factors (see Ronkowski & Iannaccone, 1989). Given the array of educational experiences and skills required for the development of a comprehensive image of leadership effectiveness, it seemed important to take a longitudinal and developmental view of administrators' socialization experiences. Thus we decided to employ this model as the second side of a two-way conceptual matrix (see Figure 1) to accommodate potential variations in the impact of the various socializing influences over time.

The Van Gennep (1960) model suggests a process of general professional maturation progressing from being defined by others to being self-defined. According to Van Gennep, at the separation stage people are concerned with comparing themselves with others and how others judge their adequacy. At the transition stage the pattern of "self location" is against the standards imposed by the functions of the job and task performance. At the incorporation stage individuals make comparisons between their former and present self (e.g., perceptions of progress made from a previous self toward becoming an instructional leader).

More recently, Leithwood et al. (1992) have described principal socialization within five dimensions; in a manner consistent with Greenfield's (1985) claims concerning mediating influences on principals' socialization (work setting, culture, and relationships). The five dimensions in the Leithwood model are relationships with superiors, relationships with peers, district policies and procedures, formal training, and outcomes. The last dimension incorporates several subdimensions: image of the role, skills, norms and values, and communication networks.

Our own framework, used as the second side of our conceptual matrix, (see Figure 1) employs four dimensions. These are internal processes (values, cognitive processes), relationships (with superiors, peers, and subordinates), organizational and contextual factors, and image of the role.

Two key differences distinguish the model employed in this study from the Leithwood framework. Conceptually we wished to recognize the influence of internal processes, such as personal and professional values (Hodgkinson, 1987), as a separate dimension. This arrangement of the framework dimensions illustrates our particular interest in assessing both the relative importance of personal values as an influence on socialization and the individual's perception of socializing experiences. Previous research (Begley & Leithwood, 1990; Campbell-Evans, 1991) suggests to us that personal values may be antecedent to other socializing influences. The Leithwood model, in contrast, focuses more particularly on a collective organizational perspective.

The second difference in our model involves the elimination of Leithwood's outcomes dimension altogether and the repositioning of several subdimensions. For example, in our framework "image of the role" is treated as a separate dimension. We hypothesized that aspiring principals have an image of the principal's role (however appropriate or inappropriate it may be) prior to, as well as during and after, the socializing experiences associated with formal training.

Other minor variations incorporated in our framework include repositioning "communication networks" as an "organizational and contextual factor," and adding "critical events" as another variation of contextual factors. The term *contextual factors* incorporates into our framework the potential influence of organizational policies and procedures on socialization as well as that broad range of potential environmental influences identified in planned change literature (Fullan, 1991; Huberman & Miles, 1984). For example, the concept of critical events is taken from the "intervention taxonomy" of Hall and Hord (1984). This subdimension permits the accommodation of perceptions reported by course candidates that they view as particularly significant or hinging influences on their socialization.

Procedures of Data Collection and Analysis

Approximately 90 candidates enrolled in two sections of the NWT Principal Certification Program implemented by OISE during 1989-1990 constituted the sample for this inquiry. Data collection for the study was conducted in three ways; a pre-course survey, interviews with individual candidates conducted by a participant observer, and access to dialogue journals kept by candidates and instructional staff.

The pre-course survey was distributed to all candidates by mail several weeks prior to the beginning of the 10-day course. Completed surveys were collected on the first day of the program. As expected, a few candidates (10 to 15) arrived for the course without the survey completed. However, a 100% return rate was achieved by requesting that these individuals complete the survey within the first two days of the program. The questions in the survey were organized in six sections; entry process, career aspirations, goals, image of the principalship, decision making, factors, and strategies. Completion of the survey provided information on why candidates enrolled in the course, what factors influenced their decision, prior leadership experiences or training, their image of an effective principal along several key dimensions (goals, factors, strategies, and decision-making), and their personal expectations for the course.

Once the course was underway, one investigator acted as course principal while the other worked as a participant observer and interviewer. Seven candidates, all aspiring principals as opposed to incumbents, were approached and subsequently agreed to be interviewed at intervals throughout the course. Several of the candidates also shared the contents of staff/candidate dialogue journals. Additional data of a corroborative nature were gathered from discussions with the instructional staff, sitting in on large and small group sessions, and from an examination of summative course evaluation data.

Preliminary analysis of the survey data was conducted by two individuals. One person, who had acted as participant observer during the course, was able to combine survey data with individual interview data. A second investigator, an OISE graduate student and experienced school administrator from the Northwest Territories, provided a second perspective that proved valuable.[3]

Two separate procedures were employed to analyze 87 pre-course surveys. One investigator used an a priori framework for classifying data (see Figure 1). He analyzed each survey along three dimensions (relationships, organizational/contextual influences, and image of the role) from the conceptual model described as Figure 1. A second investigator used a grounded approach of analysis by focusing attention on candidate responses to a subset of the questions from the surveys. These were questions relating directly to entry process and career aspirations. This second procedure focused directly on the candidates' responses to specific questions as the means of identifying the factors or themes they perceived as influential to their socialization.

Once the initial analysis of the 87 surveys was completed by the two investigators using separate procedures, the themes and influences identified by the grounded approach were matched against the detailed analysis of all survey responses produced through the conceptual framework procedure. For reporting purposes, and to facilitate further analysis, the themes identified by the grounded procedure were grouped according to the four dimensions of the conceptual framework (e.g., see Table 1). In this way it was possible to compare the socializing influences identified from direct questioning in the surveys with the results of the broader analysis of all survey responses using the conceptual framework. The degree of congruence in findings from the two investigators became quite apparent.

Results and Discussion

The findings are presented in three sections corresponding to the research questions: factors that influenced decisions to pursue the principalship, factors influencing decisions to pursue training, and factors influencing decisions to enroll in a particular training course. We also attempt to address some of the questions that naturally emerge from our theoretical framework.

Factors That Influence Decisions to Pursue the Principalship

Table 1 summarizes the data relating to decisions to pursue a principalship. Ten themes were identified through the grounded analysis of a focused subset of questions and responses from the pre-course survey. A subsequent matching of these themes against the profiles generated through a second data analysis, employing the four-dimensional conceptual framework, produced no additional themes. However, the frequency with which respondents could be linked with some themes tended to be higher with the conceptual framework analysis, presumably because these analyses were based on more data (responses from the complete survey rather than a selected subset of questions).

Personal motivations, or internal processes, appear to figure prominently as key influences on decisions to pursue the principalship, a category of influence on socialization that has been relatively unexamined in previous studies. Of the 10 themes identified and grouped within the four dimensions of the conceptual framework, it became apparent that the three most frequently identified themes (challenge, interest, and responsibility) were most appropriately grouped within the internal processes dimension. Over 30% of the 87 candidates spoke in terms of "challenge" and/or "interest" as the basis for their decision to pursue the principalship. Other strongly represented themes were distributed fairly evenly across the four dimensions of the conceptual framework. For example, 22% of the candidates suggested they were in-

Table 1
Factors That Influence Decisions to Pursue the Principalship

Factor	*Frequency (N=87)*	
	grounded analysis	*framework analysis*
Internal Processes:		
Challenge	20	30
Interest	17	26
Responsibility	15	17
Relationships:		
Nominated/Asked/Encouraged	11	14
Desire to Share and Be Helpful	6	19
Organizational/Contextual:		
Career Advancement	7	14
To Gain Broader Experience	7	6
Make a Difference	6	15
Image of the Role:		
Belief I Can Contribute	12	13
Possess Appropriate Skills	10	16

fluenced by relationships with others (nominated or encouraged by others, and a desire to share and be helpful). A similar number of candidates explained their pursuit of the principalship according to their perceptions of the role, expressed in terms of a belief that they could contribute and the extent to which they already possessed the necessary knowledge and skills.

It was interesting to note that themes grouped as organizational and contextual influences (such as career advancement, broader experience, and desire to make a difference) were less well represented. Organizational arrangements and professional relationships were apparently not the strongest incentives for attracting individuals to consider school leadership roles. Of course, it is possible to argue that the Department of Education was just not doing the right things organizationally to attract principal candidates.

Comparing these findings with those reported by Duke (1988) in his study of incumbent principals produces some interesting parallels. Duke's inquiry, focused on why principals consider quitting their profession, discovered that incumbents derive considerable satisfaction from several aspects of the job. These include the diversity of tasks, the numerous opportunities to solve complex problems, and the chance to learn more about their own abilities and beliefs. It is apparent that the reasons offered by aspirants in our study for seeking a principalship are quite similar to those aspects of the role identified in the Duke study as most satisfying to incumbents. In both studies, personal perceptions of challenge and interest figure prominently as influences. This implies that some influences on socialization may remain constant throughout Van Gennep's (1960) three hypothesized stages of development (separation, transition, and incorporation). However, none of the principals in the Duke study mentioned responsibility and advancement as key motivators for staying in the role. In our study of aspiring principals these two themes were identified as moderate influences on a decision to pursue the principalship.

Ironically, Duke discovered that certain aspects of principals' work appear to produce concern and frustration at the same time as they produce satisfaction. Although both aspirants and incumbents are apparently attracted to the principalship by the rich variety of challenges to the intellect associated with the role, this same challenge of doing all the things that principals are expected to do also motivates incumbents to consider leaving the job. Duke concludes with skepticism toward the notion that the conflicts experienced by these principals could have been reduced by more effective training and socialization. He suggests that in some cases their problems were more likely functions of their personalities than of the job. Thus the irony becomes extended by the findings of our study. The internal processes that were discovered to be the strongest influence on aspirants deciding to pursue the principalship may be the same forces that satisfy current incumbents as well as cause other veteran principals to consider leaving the role.

Factors That Influence Decisions to Pursue Training

Just as decisions to pursue the principalship are apparently not heavily influenced by relationships with others, decisions to pursue training appeared to be only modestly influenced by superordinates. This is an apparent contradiction of Van Gennep's (1960) theorized importance of modeling by others at the separation stage. Findings in our study suggest that decisions to pursue

training are again most likely to be influenced by internal processes, a dimension that was not considered in either the Greenfield or the Leithwood principal socialization studies. More specifically, having decided to pursue the principalship, more than half of our respondents apparently pursue training as a way of moving toward an existing image of the role. This appears to confirm our hypothesis that some candidates possess an image of the role that may be antecedent to other socializing influences. However, it must be emphasized that the images of the role manifested by these candidates were varied, often only partially developed, and not necessarily consistent with the notions of instructional leadership.

A total of 12 themes were identified in our study as influences on candidates' decisions to pursue training for the principalship. Ten of these themes were identified by the analysis of a focused subset of survey questions and responses, that is, those relating to entry process and career aspirations. The subsequent broader analysis of all survey data, using the four-dimension conceptual framework (Figure 1), produced an additional two themes that appeared frequently enough to be included in the list. These were "professional development" and "personal enrichment."

As Table 2 illustrates, four themes were most frequently identified by the data analysis. The themes "skill development" and "pursuit of knowledge" were cited in 49 and 42 cases respectively; approximately 50% of the 87 respondents. A third frequently identified theme was "step in the career ladder." It was cited as a motivator to pursue training in 30 surveys, or approximately 33% of candidates. A final theme identified by 26 respondents was "desire to broaden or gain perspective."

When these 12 themes were grouped within the four dimensions of the conceptual framework (see Table 2) it became apparent that themes were

Table 2
Factors That Influence Decisions to Pursue Training

Factor	*Frequency (N=87)*	
	grounded analysis	*framework analysis*
Internal Processes:		
Desire for Knowledge	15	42
Interested	11	15
Curiosity About the Role	2	7
Professional Development	-	7
Personal Enrichment	-	14
Relationships:		
Collegial Contact	1	15
Organizational/Contextual:		
Step in Career Ladder	22	30
Present Role Complements	6	8
Desire for Training	2	15
Image of the Role:		
Skill Dev./New Principal	11	49
Broaden or Gain Perspective	5	26
Relevance to Present Job	4	19

again heavily clustered in the internal processes dimension, and to a lesser extent in two other dimensions, organizational/contextual factors and image of the role. As was the case for factors influencing decisions to pursue the principalship (reported in the previous section), the relationships dimension was not a particularly significant category of influence. Only one theme, "collegial contact," was identified from 15 surveys and grouped within this dimension. On the other hand, whereas organizational/contextual factors were apparently not strong influences on the decision to seek a principalship, three themes grouped in this category were identified as influences on decisions to pursue training for this educational role. These were "step in the career ladder" (30 cases), interest in specific "training" (15 cases), and present placement in a "complementing role" (8 cases). Similarly, themes related to the image of the role dimension were much more strongly represented as influences on decisions to pursue training than on decisions to pursue the principalship itself.

The heavy clustering of themes (5 of 12) within the internal processes dimension for decisions to pursue training repeats the pattern previously established for decisions to pursue the principalship as a role. In both types of decisions, personal goals or objectives figure prominently as influences. Furthermore, if these themes are ranked within a values hierarchy ranging from personal preference to values of principle as proposed by Hodgkinson (1987) (see also Begley & Leithwood, 1990), these themes also imply ethic related, or transrational aspirations for "knowledge," "professional development," and "personal enrichment." While not conclusive indicators of high psychological maturity, these themes do suggest that the candidates are not articulating personal preference orientations in the sense of pursuing training as a path to great power or wealth. Given the lack of clarity and general ambiguity associated with much of the principal's role (Duke 1988), and the demands for self-direction and leadership associated with an instructional leadership orientation (Leithwood & Montgomery, 1986), the presence of these themes is reassuring. The NWT training course appears to be attracting the right candidates, those committed to a self-directed and professional orientation to the role.

As has been previously suggested, candidates are apparently much less influenced by their relationships with others in deciding to pursue training than might be expected. The one exception was an apparent desire in 15 cases for increased collegial contact, presumably with peers. Superordinates appear to have little influence on decisions to pursue training among this group of candidates. This is something of a paradox in that NWT entry requirements for the program include the endorsement of the candidates' supervisors of schools. Furthermore, it appears to be a direct contradiction of Van Gennep's (1960) theoretical assertion that at the separation stage of the socialization process candidates tend to compare themselves with others (Ronkowski & Iannaccone, 1989). Our data suggest that the decision to pursue the principalship and the decision to pursue training are not heavily influenced by relationships with others.

On the other hand, findings do suggest that decisions to pursue training are most likely to be influenced by internal processes. Having decided to

pursue the principalship, more than half apparently pursue training as a way of moving toward an existing image of the role. In our study more than 50% of candidates cited skill development as motivating them to pursue training. Nearly a third (26 cases) sought training to broaden their perspective, and 19 candidates saw the training as relevant to their present position. The implication for those interested in promoting principal recruitment, or encouraging individuals to pursue training, is that their best strategy may be to emphasize the organizational "career ladder" and "skill development" aspects of principal training, and the general value and applicability of the training to presently held positions. This is consistent with Van Maanen's (1984) notion of "chains of socialization," cited by Feldman (1989), which suggests that the process of organizational socialization consists of aspirants making only minor modifications in their behavior and attitudes from job setting to job setting.

Factors That Influence Decisions to Enroll in the OISE Course

Table 3 lists the themes identified as influences on the decision to enroll in the OISE developed Northwest Territories Principal Certification Course. Only four themes were identified in the initial analysis of a reduced set of survey questions. The subsequent analysis of all survey responses using the four-dimension conceptual framework produced an additional five themes.

Two themes dominated candidates' responses to these questions. Over 75% of candidates (67 cases) mentioned the receipt of financial support to attend the course as a factor in their decision to enroll in the course. Almost as many (62 cases) cited the availability of the NWT sponsored course as a positive encouragement to enroll. A third highly represented theme (38 cases) was the encouragement of others to enroll in the course. These influential

Table 3
Factors That Influence Decisions to Enroll in an OISE Course

Factor	*Frequency (N=87)*	
	grounded analysis	*framework analysis*
Relationships:		
Encouragement by Other Persons (colleagues, family, friends)	-	38
Organizational/Contextual:		
Certification Requirement	yes	16
Financial Support Available	yes	67
Prior Participant Endorses	yes	16
Availability of This Course	yes	62
Timing & Duration of Course (+ influence)	-	3
Timing & Duration of Course (– influence)	-	16
Employment Requirement	-	7
Toronto Site (+ influence)	-	1
Toronto Site (– influence)	-	8
Fees/Residence Expense (– influence)	-	7

persons included colleagues, family, and friends. Of the remaining six themes identified, only three were identified with any frequency (16 cases each). These were "certification requirement," "positive support of prior participants," and the negative aspects of "time" away from family and loss of holidays.

When the nine factors relating to enrollment in the OISE course were matched against the dimensions of the conceptual framework (Figure 1), an interesting reversal of the pattern set by the findings for the previous two research questions became apparent. As previously reported, candidates' decisions to pursue both the principalship and training were strongly linked to themes clustered within the internal processes dimension and to a lesser extent the image of the role dimension. In contrast, the decisions to enroll in a particular course developed by OISE and sponsored by the Northwest Territories Department of Education were apparently largely motivated by factors most appropriately grouped within the relationships dimension and particularly the organizational/contextual dimension.

From nine themes identified in the analysis, seven were grouped within the organizational/contextual factors dimension. Foremost as influences were the direct sponsorship and explicit endorsement of the course by the NWT Department of Education, combined with the financial support provided by the Northwest Territories Teachers' Association in the form of professional development travel funds. These were the two themes mentioned most frequently by candidates. The findings of this study suggest that financial subsidies to candidates and official sanction of the program are a winning combination of influences for those interested in promoting the training of principals in their jurisdictions.

Two additional organizational factors were apparent but less frequently identified. These were the perception that the course was a mandatory requirement for principals (7 cases) and the recent passing of territorial legislation relating to certification of principals. In 1989 legislation was passed in the Northwest Territories requiring all principals to become certificated within two years of being appointed principal. Other lesser influences mentioned by candidates were the timing and location of the course and the costs associated with fees, travel, and residence.

Although findings for the first two research questions suggest that relationships were not strong influences on decisions to either pursue a principalship or seek training, our data suggest that once a decision to pursue training has been made, the actual choice of training program can be significantly influenced by others. In the case of the Northwest Territories program, the steady encouragement provided by regional officials and incumbent principals to their subordinates (especially aboriginal educators who are currently underrepresented among school administrator ranks) appears to have had a positive impact. Similarly, the perceived quality of the first cycle of the program, as reported by the participants, appears to have positively influenced candidates to enroll in succeeding cycles of the program. Two implications may be drawn here. One is that Van Gennep's theorized orientation to others at the separation stage only became apparent in our data when candidates made decisions about selecting a training course. The influence of others on these decisions is somewhat limited otherwise. Second, a principal preservice

program perceived as being of high quality will apparently attract candidates to enroll, but only if they have decided on their own to pursue training, a decision usually based on personal motivations.

Conclusion

The findings of this study suggest that the socializing factors that most significantly influenced decisions to pursue both the principalship and training programs for that role are those motivators derived from the individuals' internal processes. For example, many aspirants are apparently attracted to the challenge of the principalship. Others report a thirst for knowledge, a commitment to acquiring professional knowledge, or a belief they can make a difference as motivators. Paradoxically, organizational/contextual factors and relationships with others appear to become important influences only much later in the process. Many incumbent principals who completed the first cycle of the Northwest Territories course in 1989 proved to be highly enthusiastic agents of recruitment for the second cycle of the course. The perception that the course and certification as a principal would soon become a territorial mandate was apparently another influence. Finally, many candidates identified the financial support provided by the teachers' federation as a significant influence on their decision to enroll in the program.

In addition to encouraging further research on the influence of internal processes in the socialization of aspiring school leaders, our hope is that this research will provide practical insights into how programs for school principals should be designed, implemented and evaluated. Further data are being collected through replications of this study using Ontario and Western Australian course candidates as subjects.

Notes

1. Two excellent examples of this trend are the recent legislation mandating a principal certification program in the Northwest Territories of Canada, and program development work currently being carried out in Western Australia for the Ministry of Education by a consortium of university faculty.
2. Several regionally specific profiles of the principalship are currently being developed by the authors: The Western Australia Principal Profile, November 1991; The Southwestern Ontario Principal Profile, April 1992; The Northwest Territories Principal Profile, May 1992. The results of these profiling processes were reported at the 1992 Annual Meeting of the American Educational Research Association in San Francisco. Copies of the profiles are available by contacting the authors directly.
3. We wish to acknowledge the contribution to data analysis made by Allan Brownridge, a graduate student in the Department of Educational Administration at the Ontario Institute for Studies in Education, Toronto, Ontario.

References

Argyris, C. (1982). *Reasoning, learning, and action: individual and organizational*. San Francisco, CA: Jossey-Bass.

Begley, P.T., & Cousins, J.B. (1990). *The pre-service preparation of school administrators: A description and assessment of the OISE program*. Paper presented at the annual meeting of the Canadian Society for Studies in Education, Victoria, British Columbia.

Begley, P.T., & Leithwood, K.A. (1990). The influence of values on school administrator practices. *Journal of Personnel Evaluation in Education, 3*, 337-352.

Brim, O.G. (1966). Socialization through the life cycle. In O.G. Brim & S. Wheeler (Eds.), *Socialization after childhood* (pp. 1-49). New York: Wiley.

Burdin, J.L. (Ed.). (1989). *School leadership*. Newbury Park: Sage.

Campbell-Evans, G.H. (1991). Nature and influence of values in principal decision making. *Alberta Journal of Educational Research, 37*(2),167-178.

Daresh, J.C., & Playko, M.A. (1992). *The professional development of school administrators*. Toronto: Allyn and Bacon.

Duke, D.L. (1987). *School leadership and instructional improvement*. New York: Random House.

Duke, D.L. (1988). Why principals consider quitting. *Phi Delta Kappan, 70*, 308-312.

Feldman, D.C. (1989). Socialization, resocialization, and training: Reframing the research agenda. In I.L. Goldstein & Associates (Eds.), *Training and development in organizations* (pp. 376-416). San Francisco: Jossey-Bass.

Fullan, M.G. (1991). *The new meaning of educational change*. Toronto: OISE Press.

Greenfield, W.D. (1985, April). *Being and becoming a principal: Responses to work contexts and socialization processes*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Greenfield, W.D. (1987). *Instructional leadership*. Boston: Allyn & Bacon.

Hall, G.E., & Hord, S.M. (1984). Analyzing what change facilitators do: the intervention taxonomy. *Knowledge: Creation, Diffusion, Utilization, 5*(3), 275-307.

Hallinger, P., & Murphy, J. (1987). Instructional leadership in the school context. In W. Greenfield (Ed.), *Instructional leadership: concerns, issues, and controversies* (pp. 179-203). Boston: Allyn and Bacon.

Heck, R.H., Larsen, T.J., & Marcoulides, O.A. (1990). Instructional leadership and school achievement: validation of a causal model. *Educational Administration Quarterly, 26*, 94-125.

Hodgkinson, C. (1987). *Towards a philosophy of administration*. Oxford: Blackwell.

Huberman, M.A., & Miles, M.B. (1984). *Innovation up close: How school improvement works*. New York: Plenum.

Leithwood, K.A. (1988). *A review of research on the school principalship*. Study commissioned by the World Bank, Washington, D.C.

Leithwood, K.A., Begley, P.T., & Cousins, J.B. (1990). The nature, causes and consequences of principals' practices: an agenda for future research. *Journal of Educational Administration, 28*(4), 5-31,

Leithwood, K.A., Begley, P.T., & Cousins, J.B. (1992). *Developing expert leadership for future schools*. London: Falmer.

Leithwood, K.A., & Montgomery, D.G. (1986). *Improving principal effectiveness: The principal profile*. Toronto: OISE Press.

Leithwood, K.A., & Steinbach, R. (1991). Indicators of transformational leadership in the everyday problem solving of school administrators. *Journal of Personnel Evaluation in Education, 4*, 221-244.

Leithwood, K.A., Steinbach, R., & Begley, P.T. (1992). The nature and contribution of socialization experiences to becoming a principal in Canada. In F.W. Parkay & G.E. Hall (Eds.), *Becoming a principal: The challenges of beginning leadership* (pp. 284-307). Toronto: Allyn & Bacon.

Ronkowski, S., & Iannaccone, L. (1989). *Socialization research in administration, graduate schools and other professions: The heuristic power of van Gennep and Becker models*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Schein, E.H. (1968). Organizational socialization and the profession of management. *Industrial Management Review, 9*, 1-16.

Van Gennep, A. (1960). *The rites of passage*. Chicago: University of Chicago Press.

Van Maanen, J. (1976). Breaking in: socialization to work. In R. Dubin (Ed.), *Handbook of work, organization, and society* (pp. 67-130). Skokie, IL: Rand-McNally.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 4, December 1992, 301-315*

Barry Onslow

Carol Beynon

and

Arthur Geddis
University of Western Ontario

Developing a Teaching Style: A Dilemma for Student Teachers

One aspect of the complex and arduous task of learning to teach is the acquisition of the pedagogical content knowledge that enables a teacher to transform particular content knowledge into forms accessible to pupils. This article describes the developing pedagogical content knowledge of two student teachers enrolled in a one-year elementary teacher education program. In particular it describes their deliberations about how to teach conversion between metric linear units and why the selection of any one teaching strategy can create confusion and anxiety for novice teachers. The student teachers' attempts to find solutions are presented in the form of several short vignettes, followed by discussion and interpretation of the observations.

Un problème particulièrement épineux dans la tâche complexe et ardue d'apprendre à enseigner est l'acquisition des connaissances du contenu pédagogique qui permettra à l'enseignant(e) de transformer les connaissances particulières du contenu en formes accessibles aux élèves. Cet article décrit le développement de l'acquisition du contenu des connaissances pédagogiques chez deux stagiaires inscrits au programme d'un an de formation de l'enseignant(e) à l'élémentaire. On décrit en particulier comment leurs délibérations progressent concernant l'enseignerment de la conversion en unité métrique linéaire et pourquoi la sélection d'une stratégie entre autres, peut rendre l'enseignant(e) débutant(e) confus(e) et anxieux(se). Les solutions proposées par les stagiaires sont présentées en forme de courtes vignettes suivies de discussions qui tentent d'interpréter les observations.

Introduction

Coping with the realities of teaching during the early survival years is anything but trivial. Student teachers are attempting to develop their own styles of teaching and concurrently satisfy both external and internal validations. Externally, student teachers have to "pass the test" (Calderhead, 1987), that is, they

Barry Onslow is an associate professor and Co-ordinator of the Mathematics Education Group. His research interests include the diagnosis of conceptual obstacles hindering students' understanding of mathematics and the strategies with which educators can better facilitate learning.
Carol Beynon is Director of Practicum Services and Co-ordinator of Primary/Junior Teacher Education. Her current research interests include the investigation of pedagogical content knowledge of novice teacher educators.
Arthur Geddis is an assistant professor in the Division of Curriculum Studies. His research interests include the investigation of pedagogical content knowledge of science teachers and teacher educators, and the role it plays in transforming subject matter content for instruction.

have to meet the expectations of both cooperating teachers and faculty advisors. Personally, student teachers have to wrestle with the various views of teaching that they have internalized. Teaching is unique among all the professions in that novices already have been intimately involved with the process as learners in the classrooms that they now enter as teachers. Lortie (1975) estimates that American students spend approximately 13,000 hours in direct contact with classroom teachers during their public school experience. During this "apprenticeship of observation" students internalize many expectations of "what teaching is" and "what teachers do." This observation continues on through their undergraduate and preservice programs.

During the early stages of practice teaching, student teachers have a tendency to develop a relatively utilitarian perspective on teaching in which the major criterion for success is "getting the class through the required lesson on time in a quiet and orderly manner" (Zeichner, 1980, p. 49). Once they progress past this stage, they are ready to risk experimenting with different teaching styles. However, it is only when student teachers are at the point where they start to ask "Why should I teach this way?" as well as "How do I teach?" that they progress from being a simple technician to what Schön (1987) calls a *reflective practitioner*. These particular questions are not simply general questions, but are embedded in the subject matter that people are engaged in teaching. Without the ability and desire to analyze their practice, teachers of mathematics, for example, often limit themselves to teaching for what Hiebert and Lefevre (1986) describe as *procedural knowledge* (familiarity with the formal language of mathematics and the memorization of rules and algorithms to solve mathematical tasks) instead of *conceptual knowledge* (a connected body of knowledge in which the linking of relationships is as significant as the discrete mathematical tasks). When student teachers who have learned mathematics through the memorization of rules and their precise manipulation are exposed to strategies that encourage teaching for conceptual knowledge, the conflict of how to teach can be extremely disconcerting.

This article describes the developing *pedagogical content knowledge* (Shulman, 1986) of two student teachers, how they decide to teach a mathematical topic, and why the selection of a particular teaching strategy can create confusion and anxiety for novice teachers.

The Study

The study focuses on a dilemma faced by two student teachers as they attempt to transform their understanding of a topic in mathematics, namely the conversion of metric linear units into a form that can be understood by their pupils. The student teachers' understanding of pedagogy at this point in their career is gradually being shaped and reshaped to fit a style of teaching with which each of them feels comfortable. Often the style of teaching advocated by university faculty is in conflict with the style of teaching remembered by student teachers during their own schooling. Many student teachers remember mathematics classrooms in which the teacher *tells* and the pupils *remember*. Taking on the role of facilitator, helping learners to construct conceptual knowledge, emerges as an exciting and rewarding alternative. This style of teaching, however, involves the transformation of subject matter knowledge,

and so requires a firm grasp of various components of pedagogical content knowledge.

Student teachers' over exposure to any one particular style of teaching can prevent a critical appraisal of alternative strategies and limit their teaching. Recognition of this predicament and the student teachers' attempts to find solutions are presented in the form of several short vignettes, followed by a discussion of the student teachers' observations.

Conceptual Framework

Traditionally, the preparation of novice teachers for classrooms has focused on two broad but distinct areas: content knowledge and general pedagogical skills. For the prospective elementary school teacher emphasis has usually been placed on general pedagogical skills rather than content. Underlying this emphasis in pedagogy is the assumption that pedagogical skills are generic, that is, they are by and large independent of the content being taught. Shulman (1986) has drawn attention to the need to abandon this simplistic assumption and to pay more attention to the ways in which pedagogy and subject matter interact. To a significant degree, the intellectual expertise of teachers can be seen to lie in their ability to *transform subject matter knowledge* into forms that are accessible to their pupils. This ability relies on a teacher's awareness of those aspects of content knowledge particularly relevant to its teachability. Shulman has called such knowledge pedagogical content knowledge.

Student teachers are seldom aware of the need to transform subject matter knowledge. For them the act of teaching involves a simple transmission of subject matter knowledge—more or less intact—to their students.[1] Shulman's concept of pedagogical content knowledge would appear to have potential both for clarifying the need to transform subject matter and for assisting reflection on how to perform the transformation.

The task of learning to teach is a complex and arduous one. Unfortunately, there is little general awareness of this fact. In his Presidential Address to the American Educational Research Association, Shulman (1986) drew the attention of both policy makers and teacher educators to the inadequacy of conceptualizing the knowledge base for teaching in terms of the dual categories of subject matter and generic teaching skills. Focusing attention on the task of readying subject matter content for teaching in terms of a transformation of that subject matter knowledge, he attempted to articulate the kinds of knowledge required to perform such a transformation.

Pedagogical Content Knowledge

Shulman (1987) has drawn attention to a type of knowledge that plays a central role in the transformation of subject matter knowledge. He represented this pedagogical content knowledge as a "special amalgam of content and pedagogy that is uniquely the province of teachers, their own special form of professional understanding" (p. 8).

> It goes beyond knowledge of subject matter per se to the dimension of subject matter knowledge *for teaching*. [It is that] particular form of content knowledge that embodies the aspects of content most germane to its teachability. (Shulman, 1986, p. 9)

At one level pedagogical content knowledge might be dismissed as simply better, more complete understanding of the subject matter. Such a move, however, would blur an important distinction that has considerable potential for teacher education. Almost every teacher has had the experience of uncovering new aspects of subject matter while engaged in teaching it. The well-known expression "You don't really understand a topic until you've taught it" gives credence to Shulman's view that understandings rooted in the intersection of subject matter and pedagogy are central to the task of learning how to teach. For novice teachers in particular, it is important to realize that there is important knowledge about content central to the task of preparing that content for teaching.

In order to be able to transform subject matter content knowledge into a form accessible to pupils, teachers need to know a variety of particulars about the content that are relevant to its teachability, particulars that probably would not have been revealed until the task of teaching the subject matter had been assumed. This pedagogical content knowledge is in some sense a result of the interaction of content and pedagogy. It is knowledge about the content that is derived from consideration of how best to teach it. For a particular topic, pedagogical content knowledge includes knowledge of:

1. what makes the topic easy or difficult to understand, including the preconceptions about the topic that pupils bring to their studies;
2. those strategies most likely to be effective in reorganizing pupils' understanding to eliminate their misconceptions;
3. a variety of effective means of representing the ideas included in the topic: analogies, illustrations, examples, explanations, and demonstrations (Shulman, 1986, pp. 9-10).

There is no reason to assume that these categories exhaust the possible kinds of pedagogical content knowledge. They provide, however, a comprehensive framework for initial investigations. Student teachers are keenly aware of their need to acquire knowledge and skills in the area of general pedagogy: for example, motivation, questioning skills, classroom management. However, they are usually surprised to discover that there are many aspects of content knowledge relevant to the task of teaching that they need to acquire. This phenomenon is especially true for student teachers enrolled in elementary teacher education. The acquisition of pedagogical content knowledge is often a crucial factor in novice teachers developing the ability to communicate their knowledge to pupils. Even more important, the awareness that teachers need to transform subject matter so as to make it accessible to pupils, shifts the focus of learning to teach from the simple acquisition of a set of classroom management and knowledge transmission skills to the more demanding but intellectually satisfying task of inquiry into how pedagogy and content knowledge interact (see Figure 1).

Conceptualizing the central intellectual task of teaching as that of transforming content knowledge into a form accessible to pupils provides student teachers with a rationale for inquiring into their own grasp of content, their pupils' prior conceptions, and plausible reasons for pupils holding these conceptions. From this perspective, the task of learning to teach becomes more than just the acquisition of a set of general pedagogical skills and strategies

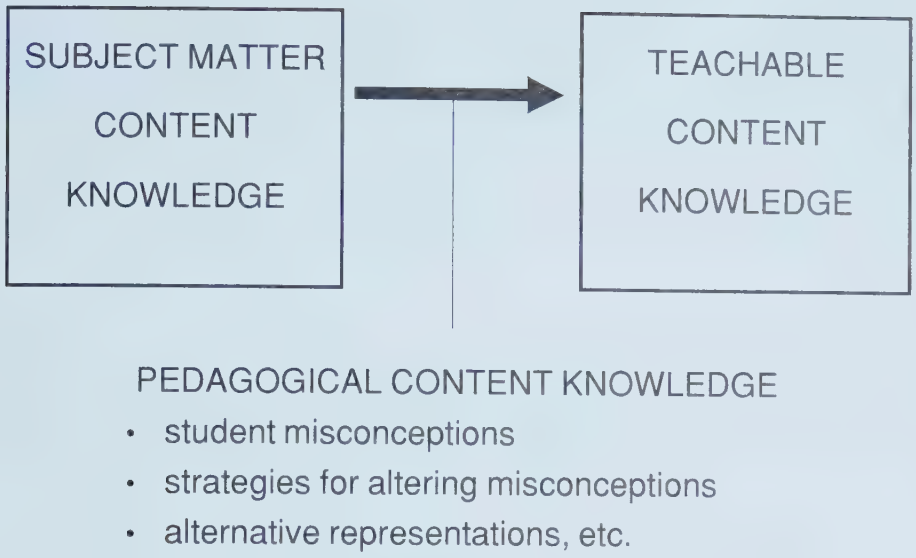


Figure 1. Transforming subject matter knowledge.

with which to transmit a fixed body of knowledge. The task of learning to teach becomes, at least in some part, an inquiry into subject matter content and its teachability. Such a stance would seem to be in tune with the reality of the complex, multifaceted, and continually changing task of the classroom teacher.

Procedural and Conceptual Knowledge

It is not possible to treat procedural knowledge and conceptual knowledge as an either/or proposition. When children learn procedures in a meaningful manner, relationships are established and the procedures are inevitably linked to conceptual knowledge. However, when mathematical procedures are learned in isolation, that is, when they are not connected to other related conceptual schema, their usefulness is severely limited. This occurs, for example, when pupils memorize the volume of a cube as $l \times l \times l$, the volume of a rectangular prism as $l \times w \times h$, and the volume of a cylinder as $\pi r^2 h$. Recognizing that all three formulae are particular cases of the more general expression: Area of base x perpendicular height, enables learners to reduce the amount of procedural knowledge that they have to store in memory, facilitates the construction of connections between the volumes of the various solids, and builds a conceptual framework that makes sense. Procedural and conceptual knowledge should inform one another, and it is an educator's understanding of the relationships between the two that assists his or her teaching of mathematics (Hiebert & Lefevre, 1986; Silver, 1986).

Methodology

The study was a naturalistic, qualitative inquiry into the developing pedagogy of two student teachers during their preservice year in an elementary teacher education program.

Name: _____

Consider an event, related to your teaching, that seems important. Do not worry about "good" language or grammar, just describe the situation as realistically as possible. If you have some second thoughts as you go along, just stick them in. Do not worry about having to stay too closely with the outline. The following categories may help you capture the details of the event.

(a) The context:*

(b) Your intentions:

(c) Who was involved:

(d) What you did:

(e) What happened as a result:

(f) How did this event make you rethink your views of teaching, learning, students, subject matter, etc.:

g) Other comments, such as expectations for the future, predictions, what was learned, and so on.

*The form consisted of both sides of an 11 x 8.5" sheet of paper with spacing after each subheading for the student teachers to summarize the critical incident.

Figure 2: Critical teaching incident sheet.

Data collected from the student teachers consisted of interviews, field notes and recordings of class teaching, and critical incident sheets. Interviews that took place prior to and following observed lessons were audiorecorded and transcribed. Further information was collected following the student teachers' submissions of critical incident sheets. These critical incident sheets were adapted from Miles' (1990) procedure for gathering data for vignettes (see Figure 2). The student teachers were requested to complete critical incident sheets for events that they considered to be significant in their professional development. The critical incident sheets were then analyzed and interviews arranged to expand and clarify the details and significance of these incidents to the students' professional growth.

The vignettes provide a means for documenting and inquiring into the knowledge growth of novice teachers (Shulman, 1987). The incidents that these vignettes capture and the subsequent analysis add to our understanding of teaching by illustrating the dilemmas faced by novice teachers as they attempt to transform content knowledge into forms that their pupils can understand.

The Student Teachers

The vignettes that follow focus primarily on Mary,[2] a confident 24 year old woman who had worked for a year after completing her honors undergraduate psychology degree prior to entering our one year elementary teacher education program. Although Mary had completed a calculus course in high school and several statistics courses in university, she readily acknowledged that she did not enjoy mathematics and felt anxious about teaching the subject. She explained that she "never remembered liking it [mathematics] or really remembering it." For Mary, math was "something that I can remember being drilled into me—like it was for a test and no other reason."

June was Mary's teaching peer.[3] She was 23 years old and had entered the faculty of education after completing her honors undergraduate degree in physical education. June enjoyed mathematics. While in high school she had successfully completed an academic mathematics program. In university, however, she had had no reason to continue her mathematics education. She was secure in her own mathematical ability, but on several occasions indicated that she found it difficult to understand why some children were unable to grasp rudimentary mathematical concepts. June felt that "coming down to the children's level" would be her most difficult task in teaching mathematics to elementary schoolchildren.

The Setting

Mary and June had been placed in the same grade 5 classroom for their three-week practicum. This session was their second practicum experience. The first had been a four-week block in a grade 2 classroom. The mathematics unit, which they were to complete over the three week period, was measurement. They had spent considerable time planning the unit and had organized it so that the children would be engaged during the first week at a variety of self-directed activity centers involving linear measurement. Instructions for each center were written on cards and the pupils worked in heterogeneous ability groups, attempting to solve the problems cooperatively. While the children worked at the activity centers, the student teachers acted as facilitators, interacting with individuals or groups as the need arose. For the final two weeks of the unit, Mary and June had planned to follow a more traditional approach. The mathematics lessons would continue to embrace the notion of linking conceptual and procedural knowledge, but would be more teacher-directed than the learning centers had been. Topics to be included were "conversion between metric linear units," and the "concepts of area and volume."

The Dilemma

Mary and June's dilemma first surfaced after their pupils had completed the self-directed activity centers on linear measurement and June had taught the first teacher-directed lesson dealing with metric conversion. On entering the practice teaching classroom during a lunch break, their faculty mathematics instructor found Mary and June at the chalkboard engaged intently in a discussion with their cooperating teacher. They were all concerned with the apparent inability of the pupils to convert from one metric linear unit to another. On the worksheet that June had prepared, involving elementary conversions such as 5m=__cm, the pupils had made numerous errors. From one point of view this should not be surprising. As Shulman (1989) has pointed out, knowing grade 5 subject matter is not "trivially simple" (p.9). It was necessary for Mary and June to comprehend their pupils' learning processes, the typical understanding of their grade 5 pupils, common errors their pupils might have acquired, and concepts that are difficult or easy for grade 5 pupils (Marks, 1990). Mary and June also required a repertoire of pedagogical representations (metaphors, analogies, examples, stories, demonstrations) that could assist pupils in connecting what they already knew to what they were about to learn (Shulman, 1989), and consequently help the pupils construct their own representations for the mathematical abstractions.

Mary had to teach the next lesson on metric conversion and was unsure how she should proceed. If she simply presented the pupils with a list of procedural rules and drilled them in the use of these rules, she was afraid that she would be promoting the same negative reaction towards mathematics that she had herself developed in elementary school. However, she was becoming increasingly frustrated with the results of the approach advocated by her faculty mathematics instructor. As Mary wrote later:

> I found that it was difficult to compromise my new views about presenting mathematical concepts with my old memories about how I learned the same material. Perhaps I was surprised and even a little annoyed at how much my own education may influence how I now educate (or attempt to educate!) others. I intended to let or allow students to discover or develop the concept of metric conversions on their own as much as possible. What interfered with this intent were time constraints and the relative difficulty of this approach as opposed to simply presenting the "rules" for students to follow and apply.

The difficulties inherent in teaching mathematics meaningfully, the time constraints involved in covering the content of the curriculum, and the need to cope with individual differences are all factors that are not unique to Mary's situation. Mary wondered how it was possible to teach for conceptual understanding within the everyday constraints of teaching. Mary's attempt to manage her dilemma (Lampert, 1985) and her reflections on that process are now discussed.

Searching for a Solution

Two questions seemed central to dealing with Mary's difficulties:

1. How was she to manage her internal conflict?
2. How should she plan to teach metric conversion effectively?

How was Mary to Manage her Internal Conflict?

Should Mary select a style of teaching that epitomizes her unpleasant memories of mathematics but that she thinks might be effective, or a style that she agrees with in theory but that seems to be proving ineffective?

Initially Mary had rejected the method by which she remembered being taught, that is, drill and rote memorization, in favor of a more meaningful approach. Two elements had led to this decision. Mary felt that her own mathematics education had resulted in her dislike and anxiety toward mathematics. However, of greater significance was the emphasis in her teacher education program on teaching for conceptual understanding. Consistent with this emphasis had been a focus on the need for children to be active participants in the learning process rather than passive receivers of knowledge.

> I don't know if it's the faculty or it's just the way ... you just hear it and hear it and hear it. Like it's not just in math, you hear it in English. You know, everything's child centered and active and manipulation.... So I think you say, well, this is the way it is now.

Mary's experiences during her first practicum had supported the message she had been receiving at the faculty. Being an active participant rather than a passive receiver of knowledge enhanced learning.

> I mean it's not that I don't agree with it [the faculty approach], and I'm just accepting it blindly. I mean, when you go out there [to the schools] you really see that it's having a great effect on how kids are learning.

Prior to the lesson in question, Mary's new belief in the importance of active learning and the development of conceptual knowledge had prevented any consideration of using the strategies employed during her own schooling. Mary's frustration with her pupils' inability to convert from one metric unit to another, however, forced her to search for alternatives. Yet the mere consideration of teaching in a manner similar to the way in which she had been taught disturbed her.

> At one point, I actually found myself defensive about the fact that I actually thought about teaching conversion in a way very similar to the way I was taught.

At first one might link this phenomenon to that of children who refuse to listen to the same music or dress in a similar fashion to their parents. However, Mary's rationale for not teaching as she had been taught was carefully considered. She had not really understood mathematics. She had memorized rules to give back on a test. Now she felt anxious about mathematics because she had forgotten many of these rules, and the subject involved only unpleasant memories for her.

After sharing her thoughts with her peer, cooperating teacher, and faculty advisor, and reflecting more deeply on her own school experiences, Mary came to see her own school mathematical experience in a different light. While she had developed some negative attitudes toward mathematics and often did not understand why certain mathematical procedures worked, she could usually execute those procedures accurately. In fact, after some review her old proficiency returned. Clearly, her own mathematics education had not been a complete failure. As a result she no longer felt compelled to discard all the strategies that had been used to teach her mathematics.

> You know, there are benefits to the way we've been taught, because we're not stupid.... I think that as today's educators, we have to (or at least I have to) accept the fact that our [own previous] education will affect our teaching from time to time, often without us realizing it. As long as I am willing to admit this, I think I can use my education experiences to better my teaching practices. Just because I have been educated as a teacher NOW does not mean my experiences as a student, good or bad, are erased.

In order to reflect critically on her own teaching, Mary had to call into question not only her school experiences, but also the approach being advocated in her teacher education program. Coming to this realization made it possible for Mary to assess her alternatives. She realized that the extreme guilt she had felt for employing a strategy used by one of her former teachers was unwarranted. Also, she realized that she need not accept blindly the approaches advocated by the faculty. Instead, what she needed to do was assess carefully the strengths and weaknesses of various possible approaches and then examine how they might be adapted to meet the specific needs of her pupils. She was becoming aware that it is only through the integration of

knowledge from many different sources that teachers can begin to plan an effective curriculum.

How Could Mary Plan to Teach Metric Conversion Effectively?

Prior to the realization of Mary's conflicting feelings, she and her peer June had developed a number of metric measurement centers and facilitated their pupils' progress through these centers. June had then taught a more traditional lesson on conversion using a worksheet to assess the pupils' understanding.

During the noon-hour break following June's lesson, both student teachers expressed dissatisfaction with the level of their pupils' understanding. The children were having difficulty and most were confused. June herself was confident in her ability to convert from one metric unit to another. She could multiply or divide by the appropriate power of 10 and move the decimal point accordingly. The strategy had little meaning for her pupils, however, and June had been unable to develop an alternative strategy during the lesson.

> I did the first few and then let them try it [the worksheet], and they didn't understand it. So I tried to re-explain it using some of the numbers on the sheet to show them ... moving the decimal and you know, you're not supposed to move the decimal point,[4] but on the chalkboard it's easier to just move the decimal.... I couldn't think of any other ways to do it. You know, I explained in different words and I changed the numbers.

Mary explained that June had spent "most of the 30-40 minutes talking to them [the children] and they just didn't get it ... after June's lesson, she was confusing herself talking!"

In an attempt to assist the children's procedural understanding Mary developed the chart below to use during the next lesson. The function of the chart was to help students determine which multiple of 10 to use and whether to multiply or divide when converting units. For example, when pupils had to convert from millimeters to meters, they would count the number of *teeth* on the chart between the given units (in this case three) to determine the power of 10 (10^3=1,000). Then, because they were moving up the chart, they would

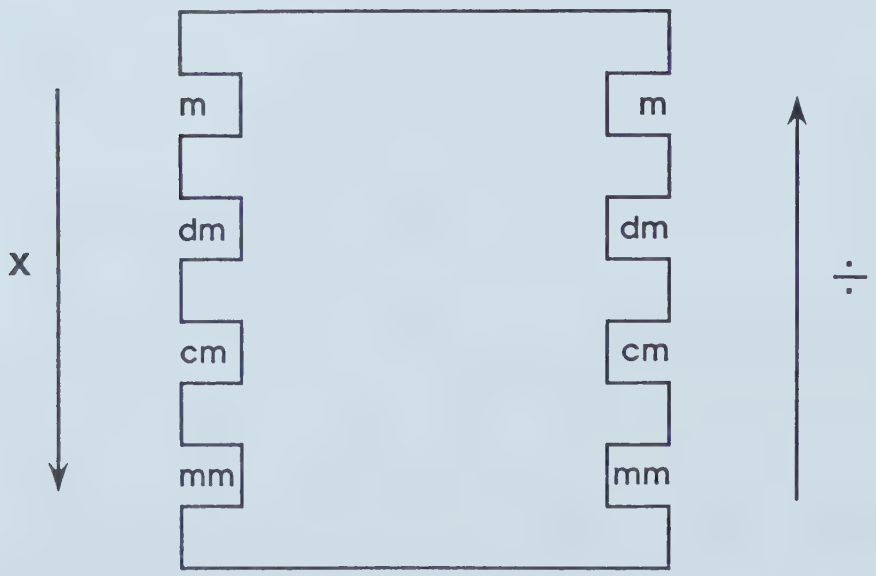


Figure 3. Mary's tooth conversion chart.

divide their original number of millimeters by 1,000 to obtain the correct number of meters. Similarly, when pupils had to convert from meters to centimeters, they would count down two teeth and so would multiply by 10^2 or 100.

Mary remembered being given a similar chart at school that had enabled her to convert metric units easily and correctly even though it was rule based. However, she was also cognizant of the need to provide pupils with more than just rules if they were to understand the procedure.

> I just didn't want to put up the chart.... I think it relieves the students of the necessity to think about what they are doing.

Mary was attempting to provide an experience that linked procedural and conceptual knowledge, while completing her lessons in what her associate teacher considered to be a reasonable period of time. Later, when Mary was asked about the usefulness of the chart, she indicated that she felt it had served its purpose, but that it had been a slow process.

> It took the full 40-minute period to go through it, and I just took that 40-minute period because I felt it was important.... It seemed to be a very slow process and I seemed at times to be going over it maybe too much. But I really wanted them to get certain points about it ... that it was not simply a rule, it was something to guide them, and to bring out the patterns of what they were really doing.

It is interesting that although Mary was the peer with the least confidence in her mathematical understanding, it was Mary who articulated what appeared to be the major conceptual obstacle faced by the pupils:

> like you're going from a smaller number to a larger number but from a larger unit to smaller unit.

Mary recognized that the pupils had a tendency to multiply when they should have divided, and divide when they should have multiplied. In converting 700 centimeters to meters, several pupils wanted to multiply because as they said, "Meters are larger than centimeters." Knowledge of the conceptual obstacles hindering children's understanding is an important component of pedagogical content knowledge. Unless teachers are aware of such obstacles they are unlikely to recognize what makes the understanding of mathematics difficult. However, while being cognizant of a pupil's difficulty is a necessary prerequisite to finding an appropriate teaching strategy, the identification of the difficulty does not automatically yield a teaching strategy that provides the desired understanding.

Mary was disappointed that the pupils' experiences at the measuring centers had not provided them with the conceptual base for understanding the symbolic manipulation.

Mary: They were all converting in the centers before we ever mentioned the word conversion, but they didn't know it.

Interviewer: Why do you think their work at the centers didn't help them with the worksheet?

Mary: I think they were understanding it from the centers, but they didn't understand they were doing conversion.... We had a couple of things where we asked for things in centimeters,

decimeters ... and they would just measure it separately each time.... I mean in a sense that's conversion.... It's not just saying, okay measure it in millimeters and then without the measuring instrument do it in centimeters and decimeters.

Interviewer: What if you had said to them, "Measure this in centimeters" then said, "Tell me what would it be in millimeters?" Would they have been completely lost?

Mary: I don't think so because a lot of them were doing that anyway.

Interviewer: But they didn't relate it to the symbolic work.

Mary: No, they didn't seem to. I didn't grasp any evidence of it anyway.

Interviewer: Perhaps that was part of the difficulty—that they couldn't relate the worksheet to what they'd been doing at the centers.

Mary: Yes. And I wonder now ... our worksheet was simply ... so many centimeters equals so many millimeters. And then the next line, so many decimeters equals so many.... If we just used a sheet that we had used at the centers ... whether that would have been it, because they would have. Oh yes, okay! I know I can do this.

Relating the symbolism of the worksheet to the pupils' previous experiences is crucial if pupils are to establish suitable representations for mathematical abstractions. Mary was beginning to understand how she could make the symbolic work more meaningful.

During the interview subsequent to the lesson, June had mentioned "baking" as a possible source of representations to assist the pupils' understanding of metric conversion. After some deliberation, however, she was still dissatisfied with the manner in which it illuminated the underlying concepts involved in metric conversion. On the other hand, when her faculty mathematics instructor suggested using money, June had a flood of ideas.

> I think that would probably be good because you could use the pennies, and the dimes and the loonies [one dollar coins], and, you know, $10 bills if you really wanted.... Monopoly money I think I would use. Yes, that's good. That would probably be good to help them. 'Cause I know they can all do it, and I never even thought of that one [money].

Because of the pressure of time Mary and June were not able to develop their new ideas during this practicum. However, they had engaged in serious reflection about how to teach metric conversion, which would no doubt influence their subsequent teaching.

Making Connections

Student teachers do not necessarily understand how to help pupils make connections between abstract symbols and the concrete learning experiences provided by activity centers. An early teaching experience of one of the authors provides a vivid example of this predicament. He remembers being introduced to Dienes' base 10 materials and thinking they would solve many of the difficulties his pupils had been having understanding the regrouping process associated with addition and subtraction. His grade 3 pupils were very happy and successful adding and subtracting with the base 10 materials, yet performed abysmally when they moved to similar paper-and-pencil cal-

culations. At the time he thought he had a poor class. Later, he realized that his pupils had an inexperienced teacher who had been unable to recognize the importance of transforming his knowledge into a form the pupils could understand.

This realization of his own limitations had occurred a few months later while attending a mathematics workshop. When he challenged the presenter on the benefit of using concrete materials, she asked him if he was helping the children move from the materials to the paper and pencil calculations. He had assumed the transition to be obvious, but why should pupils implicitly understand that the numeral 2, written in the 10s column, is equivalent to two of Dienes' 10 rods? The pupils, of course, had been performing two discrete acts (one with the materials, the other with pencil and paper) and did not recognize the connection between the two. The connection had to be made explicitly by having the pupils record and discuss the abstract symbolism while working with the concrete materials. Once the connection had been made apparent, the symbolism became comprehensible and the children used the materials as useful representations in facilitating their understanding of addition and subtraction.

It is the responsibility of faculty instructors and cooperating teachers to help student teachers develop suitable representations for the mathematics they will teach and offer guidance in bridging the gap between concrete materials and the symbols they represent.

Children have the potential to understand mathematics if given opportunities to develop connections between symbolic mathematics and appropriate representations. When pupils cannot provide a representation for an abstract mathematical concept, teachers need to be wary of the pupils' true understanding and give them the time to explore the concept further. If teachers cannot provide the representations that elucidate mathematical abstractions, it is their responsibility to search out such representations, whenever they exist, and use them to assist understanding. When one cannot provide a representation for mathematical symbolism then, generally, one does not truly understand what the symbolism represents (Onslow, 1991).

With guidance, Mary and June were able to find suitable representations that had potential for transforming their content knowledge into a form that would make it accessible to their pupils. However, without support, it is not surprising that new teachers often become frustrated with this difficult aspect of teaching.

Conclusion

Can faculty advisors and cooperating teachers help student teachers construct their own ways of knowing? Being themselves rather than cloning their supervisors is a difficult task for new student teachers. It is only with understanding and support that they will be able to meet the challenge successfully. No one correct teaching strategy is suitable for every teacher or every pupil. Many experienced teachers adopt an eclectic approach, drawing on various philosophical paradigms. Rather than being dichotomous, teaching strategies should be part of a mosaic where the pieces blend into one another to meet the many and varied situations of classroom life. Becoming knowledgeable about and comfortable with this aspect of pedagogy was an important insight in

Mary and June's development. Their experience led them to realize that it is unwise to reject any one particular style completely. "A larger and more flexible vision need not result in a rejection of traditional or familiar ideas and practices" (Feiman-Nemser & Buchmann, 1985, p. 63).

When student teachers are confronted with their faculty instructor's and cooperating teacher's beliefs and expectations over a short time span, there is little opportunity for reflection and discussion. New ideas can be accepted quickly and with enthusiasm. Yet these new ideas will probably be dismissed with similar speed if they do not produce the expected results. Student teachers need time and support to understand the intricacies of new ideas. Transforming content into a form that makes it accessible to children is anything but trivial. Developing pedagogical content knowledge involves many years of experimentation and reflection. In the reality of schools, teachers typically perform this task in isolation (Lortie, 1975). Consequently, it is important that teacher education provide student teachers with opportunities to practice their craft in a self-critical yet supportive environment where they can learn to recognize for themselves the complexities of the task and begin to learn how to deliberate on the various alternatives.

Student teachers have to meet the expectations of their faculty advisors who supervise them throughout the year, and cooperating teachers who supervise their individual practicum sessions. Receiving a poor teaching report is likely to limit a student teacher's prospects for employment. Consequently, it is naive to ignore the influence these people have on a student teacher's decision of how a particular lesson should be taught. Student teachers, like everybody else, want to be successful. When we consider the background and educational experiences of our student teachers, we should not be surprised that they are uncomfortable with the task of attempting to link conceptual and procedural knowledge. Helping pupils make connections between mathematical symbolism and suitable concrete representations is difficult to accomplish.

The attitude of faculty and cooperating teachers will greatly influence a student teacher's development. If the practicum is one of reflective inquiry in which all participants are searching rather than telling, and recognizing the many facets of teaching rather than holding one perspective sacrosanct, the practicum can be a beneficial experience. If student teachers perceive their role as conforming to one particular style, however, little will be achieved. Reflective practice in which student teachers critically analyze what they are teaching and why they are teaching it enables the student teachers to develop professionally rather than simply technically. Recognizing the important yet difficult task of assisting children in constructing representations for abstract symbolism is an essential aspect in the evolutionary development of a novice teacher's pedagogical content knowledge.

We have attempted to tell Mary and June's story using their direct quotes but with our own interpretations of their words and actions (Hall & Grant, 1991; Gudmundsdottir, 1991). Different readers will no doubt have different interpretations of the events and of our analysis. If one truly believes that every learner constructs his or her knowledge and reads each case through a different lens, then this aspect is a strength of case study research.

Notes

1. Considerable research has demonstrated that this transmission mode of teaching has been and continues to be the dominant mode of classroom instruction (see, e.g., Gallagher & Tobin, 1987; Kilbourn & Roberts, 1984; NCTM, 1991).
2. All names are pseudonyms.
3. Mary and June were part of a special project and became exceptional student teachers. They were paired with one another for each of their first two practica. This peer support provided the opportunity for student teachers to observe one another and discuss their rationale for teaching a topic in a particular fashion. The peer provided another dimension for Schön's (1987) *reflective conversation with the situation* (p. 42).
4. In the students' mathematics methods class, their faculty instructor had explained why moving the digits rather than the decimal point is a more logical process and encourages understanding rather than rote memorization.

References

Calderhead, J. (1987). The quality of reflection in student teachers' professional learning. *European Journal of Teacher Education, 10*(3), 269-278.

Feiman-Nemser, S., & Buchmann, M. (1985). Pitfalls of experience in teacher preparation. *Teachers College Record, 87*(1), 53-65.

Gallagher, J.J., & Tobin, K. (1987). Teacher management and student engagement in high school science. *Science Education, 71*(4), 535-555.

Gudmundsdottir, S. (1991). Ways of seeing are ways of knowing. The pedagogical content knowledge of an expert English teacher. *Journal of Curriculum Studies, 23*, 409-421.

Hall, S., & Grant, G.E. (1991). On what is known and seen: a conversation with a research participant. *Journal of Curriculum Studies, 23*, 423-428.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1-27) Hillsdale, NJ: Erlbaum.

Kilbourn, B., & Roberts, D.A. (1984). *Science, society, and the non-academic student: Phase I. Final Report*. Ottawa: Social Science and Humanities Research Council.

Lampert, M. (1985). How do teachers manage to teach? Perspectives on problems in practice. *Harvard Educational Review, 55*, 178-194.

Lortie, D.C. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.

Miles, M.B. (1990). New methods for qualitative data collection and analysis: Vignettes and prestructured cases. *Qualitative Studies in Education, 3*(1), 37-51.

Marks, R. (1990). Pedagogical content knowledge: From a mathematical case to a modified conception. *Journal of Teacher Education, 41*(3), 3-11.

National Council of Teachers of Mathematics, Commission on Teaching Standards for School Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

Onslow, B.A. (1991). Linking reality and symbolism: A primary function of mathematics education. *For the Learning of Mathematics, 11*(1), 33-37.

Schön, D. (1987). *Educating the reflective practitioner*. San Francisco, CA: Jossey-Bass.

Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*, 4-14.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1-22.

Shulman, L. (1989, June). Toward a pedagogy of substance. *American Association of Higher Education Bulletin*, pp. 8-13.

Silver, E.A. (1986). Using conceptual and procedural knowledge: A focus on relationships. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 181-198). Hillsdale, NJ: Erlbaum.

Zeichner, K.M. (1980). Myths and realities: Field-based experiences in preservice teacher education. *Journal of Teacher Education, 31*(6), 45-55.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 4, December 1992, 317-334*

Halia C. Silins
Flinders University of South Australia

Effective Leadership for School Reform

This study examined the relationship between school leadership and school improvement outcomes within the conceptual framework of the transformational and transactional leadership model advanced by Bass (1985; Bass & Avolio, 1988, 1990). Multivariate analyses were applied to survey data obtained from 679 teachers representing 256 elementary schools in a school improvement program. A two-factor model was supported in which transformational leadership proved to be the more dominant construct. Transformational leadership influenced all school outcomes except teacher effects. Transactional leadership influenced teacher effects favorably and school effects unfavorably. The implications of these findings in terms of leader and teacher power and their contradictory effects on school outcomes are discussed.

Cette recherche examine la relation entre le leadership d'une école et les résultats de l'amélioration scolaire en utilisant le cadre conceptuel d'un modèle de leadership transformationnel et transactionnel proposé par Bass (1985; Bass & Avolio, 1988, 1990). Des analyses statistiques multivariées ont été effectuées sur des données de sondages obtenues de 679 enseignant(e)s représentant 256 écoles primaires qui participaient à un programme d'amélioration scolaire. Les résultats indiquent une préférence du modèle à double aspect où le leadership transformationnel prédominait. Le leadership transformationnel influencait tous les résultats scolaires sauf celui d'effets sur les enseignant(e)s. Tandis que le leadership transactionnel influencait favorablement les effets sur les enseignants, il influencait défavorablement les effets scolaires. Les implications de ces résultats en fonction de leader et de pouvoir d'enseignant(e) ainsi que de leurs effets contradictoires sur les résultats scolaires y sont discutées.

Studies of exemplary schools have revived interest in the contribution made by effective leadership to school reform. Beare, Caldwell, and Millikan (1989) boldly asserted that "outstanding leadership has invariably emerged as a key characteristic of outstanding schools. There can no longer be doubt that those seeking quality in education must ensure its presence and that the development of potential leaders must be given high priority" (p. 99). However, the characteristics of outstanding leadership in schools are less obvious.

School leadership emerges from a variety of sources; it is not the prerogative of the principal. However, the principalship has been recognized as a critical element in school improvement and reform (Fullan, 1991; Sashkin, 1988). Acknowledgment of the impact of principals' leadership behaviors on school outcomes has generated an extensive body of research over the past decade (Leithwood, Begley, & Cousins, 1991). However, the thorough review of empirically based studies on the principalship conducted by Leithwood and his colleagues (1991) has revealed significant limitations to our current knowledge. This review identified three major areas of principals' influence

Halia Silins is a senior lecturer in the School of Education. Her interests and most recent publications are in the areas of leadership and school reform.

that need further empirical investigation: (a) principals' impact on a wider set of student outcomes than has been studied to date; (b) principals' impact on teacher outcomes associated with their growth and development; and (c) principals' influence on the school environment as a whole. This study addressed principals' influence on these areas using teachers' perceptions as a measure of school change.

Conceptual Framework

The transformational and transactional leadership model (Bass, 1985; Bass & Avolio, 1988, 1990) offers a range of leader behaviors that have been shown to promote change and desired outcomes outside of educational settings (Bass, 1985; Waldman, Bass, & Einstein, 1987; Yammarino & Bass, 1990). The proponent of this model, Bass, has advocated transformational leadership for successful organizational change and improved performance.

The transformational leader is one who motivates followers to perform above expectation. This transformation of performance expectation, based on a level of confidence in one's ability to achieve designated outcomes, can come about in any one of three interrelated ways (Bass, 1985; Bass & Avolio, 1988):

1. raising followers' levels of awareness regarding importance and value of designated outcomes and ways of achieving them;
2. influencing followers to transcend their own self-interest for the sake of the team, organization, or some greater mission;
3. altering followers' levels of needs or expanding followers' existing needs and wants.

Three factors determine the behavioral components of transformational leadership and define it (Bass, 1985; Waldman et al., 1987).

Charisma/inspiration. This is the degree to which the leader creates enthusiasm in followers, sees what is really important, and transmits a sense of mission to the organization. The leader inspires loyalty and devotion, instills pride and faith, and commands respect. Followers place a great deal of trust and confidence in the leader's vision and values, develop intense feelings about the leader, perceive the leader as a role model, and want to identify with him or her (Bass, 1985).

Intellectual stimulation. This is the degree to which the leader provides intellectual and problem-oriented guidance. The leader arouses followers to think in new ways (Bass, Waldman, Avolio, & Bebb, 1987). Followers are encouraged to question their own and others' assumptions, beliefs and values, and develop independent problem solving capabilities.

Individualized consideration. This is the degree to which the leader is concerned with the individual needs of followers (mentoring). The leader responds to individual differences in followers' needs for growth and development, elevating needs and abilities to higher levels when appropriate, and delegating projects to stimulate individual learning experiences. Followers are provided with coaching and teaching as needed.

The transactional leader is one who motivates followers to perform at their levels of expectation and to achieve satisfaction of basic needs. Transactions are at the heart of the interchange between leader and followers. Followers' efforts depend not only on their levels of confidence or expectations that an

outcome can be achieved through their performance, but also on the extent to which their self-interest is served. The transactional leader (Bass, 1985):

1. recognizes followers' needs and wants and establishes appropriate contingencies for performance;
2. exchanges rewards and promises of reward for followers' efforts;
3. is responsive to a follower's immediate self-interest if it can be met through getting the work done.

Two factors determine the behavioral components of transactional leadership and define it (Bass, 1985; Waldman et al., 1987).

Contingent reward. This is the degree to which the leader makes clear what the follower must accomplish in order to be rewarded. The leader provides rewards if followers perform in accordance with contracts or expend the necessary effort to meet performance standards. Clarification of goals, work standards, and assignments is emphasized. Leaders recognize what subordinates need and, through extrinsic rewards, energize followers to reach objectives.

Management-by-exception. This is the degree to which the leader provides negative feedback for failure to meet agreed-upon standards. The leader avoids giving directions if the old ways are working and allows followers to continue doing their jobs as always if performance goals are met. This is the less active form of transactional leadership. The leader intervenes only when standards are not being met.

Bass (1985) contended that most leaders exhibit both transformational and transactional leadership, in varying degrees. Transformational leadership augments transactional leadership by focusing on the development of followers as well as addressing the goals of the leader, follower, group, and organization (Bass & Avolio, 1990). For Bass, the success of a transformational leader is demonstrated both by increased performance outcomes and the degree to which followers have developed their own leadership potential and skills.

Avolio and Bass (1988) asserted that, unlike transactional leaders who work within the existing organizational system, transformational leaders change the system to recreate their environment. Transformational leaders come into their own in times of growth, crisis, and reform. Burns (1978) described reform leadership as exacting and, by its very nature, transformational. Transformational leaders tend to separate from the environment and create change (Zaleznik, 1977), whereas transactional leaders tend to work within the system to preserve the status quo.

Purpose of Study

Bass's model of transformational and transactional leadership has a number of implications for the current reform movement in education. According to the model, principals exhibiting transformational leadership behaviors will be more effective in bringing about desired outcomes, such as faculty development, improved teaching and learning, collaborative decision making, and responsive and innovative environments, than those exhibiting predominantly transactional leadership behaviors. Schools with predominantly transformational leadership are expected to be purposeful and collaborative, with a greater number of staff and faculty operating in an empowered and leaderlike

way, than those with predominantly transactional leadership (Bass et al., 1987).

Arguments linking school improvement with transformational leadership are compelling but lacking in empirical support (Hoover, Petrosko, & Schulz, 1991; King, 1989). The few existing studies (Hoover, 1988; King, 1989; Kirby, 1991) that have applied Bass's model to school settings have used perceptions of leadership effectiveness and follower satisfaction as measures of leadership outcomes. This study took the important step of addressing the impact of Bass's model on broader school outcomes within the context of a school improvement effort.

Bass's model was tested in two ways: (a) whether there is evidence for Bass's two constructs when considered in isolation using only the teachers' perceptions of leadership characteristics in their school; and (b) whether there is evidence for Bass's two constructs when examined in operation as explanatory constructs to influence the effects of leadership on teachers' perceptions of the working of the school as well as leadership characteristics. Both analyses are necessary to support the validity of Bass' model since the former considers one aspect of the validity of the model, namely the extent to which the characteristics of the leadership of principals group to form the two constructs, while the latter not only considers this aspect but also examines its utility in operation.

Survey Data

The British Columbia Ministry of Education introduced in 1989 a wide-ranging school reform policy that focused on the first three years of schooling. The policy, entitled The Primary Program, was developed in response to recommendations of the Sullivan Royal Commission Report (Royal Commission, 1987-1988). This school improvement project, referred to as The Primary Project, had in 1991 concluded its first year and began the second year of implementation of this policy. Survey data measuring perceptions of school personnel on a number of aspects of The Primary Project were collected by the Center for Principal Development, Ontario Institute for Studies in Education, University of Toronto. Access to this data base enabled this research to be carried out.

Method

Variables in Study

The two latent constructs of transformational leadership and transactional leadership were viewed as independent variables and were hypothesized to be represented by observable characteristics. Transformational leadership was hypothesized to be represented by three variables: charisma/inspiration, intellectual stimulation, and individual consideration. Transactional leadership was hypothesized to be represented by two variables: contingent reward and management-by-exception.

Four outcomes were chosen to represent the school improvement dependent variables. These were school effects, teacher effects, program and instruction effects, and student effects.

School effects. This variable encompasses perceived changes from the school improvement process impacting on the school as a whole and related to the functioning, climate and direction of the school.

Teacher effects. This variable encompasses perceived changes from the school improvement process impacting on teachers.

Program and instruction effects. This variable encompasses perceived changes from the school improvement process impacting on school programs and instruction.

Student effects. This variable encompasses perceived changes from the school improvement process impacting on students.

It was unclear whether these observable outcomes were associated with one or more than one latent construct associated with school improvement.

Operational Definition of Variables

The dependent and independent variables were operationally defined by a process of clustering items from the survey to reflect their conceptual definitions. Items from the survey were categorized by Bass and the author into subscales representing the three hypothesized behavioral components of the transformational scale and the two hypothesized behavioral components of the transactional scale. Ambiguous items were excluded. Decisions made on the basis of Bass's model resulted in the clustering of 13 items for charisma/inspiration, three items for intellectual stimulation, six items for individual consideration, three items for contingent reward, and three items for management-by-exception.

The survey also provided items that operationally defined the four scales measuring outcomes of the improvement process (school effects, teacher effects, program and instruction effects, student effects). To verify the validity of the items forming these clusters, five experienced school personnel were given definitions of the four outcome variables and asked to categorize every item in this section into one of these four scales. Only those items were used on which there was 80% agreement. Items that were difficult to categorize were excluded. Five items were identified as representing school effects, eight items for teacher effects, five items for program and instruction effects, and 10 items for student effects. The operational definitions of the variables represented by the items selected from the survey and their conceptual definitions are listed in Table 4.

Sample

Schools to be surveyed were stratified by size and randomly selected. An estimate was made of the number of elementary staff in each school and packages of the Primary Program Survey were sent out requesting that the completed survey be returned by mail. A total of 2,547 questionnaires were mailed to schools; the individual response rate was 30% and the school rate was 67%. The low response rate for teachers was at least partly due to considerable political instability in the province at the time of the survey as well as related teacher industrial action in many districts. Responses from 770 individual staff were returned: the 74 responses from principals were automatically excluded from this study and a further 17 responses were not usable because the respondents did not identify their position in the school. A sample

of 679 individual responses representing 256 schools remained. This provided a more than adequate cases-to-variables ratio for the analyses undertaken. Table 1 summarizes the population, the selected, and the achieved sample by strata.

The sample employed was a probability proportional size random sample of school districts with a sampling fraction of approximately one third. In a second stage, schools were sampled from within the chosen districts. Because a cluster sample of schools within districts was employed, and not a simple random sample of schools, the number of schools responding cannot appropriately be used in the significance tests applied to the analyses carried out. As a consequence, it was necessary to examine the *F* values obtained and their levels of significance with caution. The data obtained from the individual teachers within schools were aggregated to the school level in analysis.

Source of Leadership

Of responders 73.4% included the principal as at least one of those providing leadership in the school regarding the Primary Program project; this included 35.2% who identified the principal as the sole provider of leadership. Thus 26.6% did not identify the principal as providing leadership outside of leadership teams and committees. This supported the view of the principalship as a major source of leadership contributing to the school improvement process, although not always the sole source.

Research Design

This study employed a correlational design using factor analysis and canonical analysis models to determine the relationships between transformational and transactional leadership behaviors, and school improvement outcomes. First, confirmatory factor analysis was employed to test the two-construct model. Then canonical correlation was used to examine the relationship between the set of dependent variables (school effects, teacher effects, program and instruction effects, student effects) and the independent variables (charisma/inspiration, intellectual stimulation, individual consideration, contingent reward, management-by-exception).

Results

The logical allocation of items to scales on the basis of Bass's model of leadership was tested against the interrelations of items obtained empirically. Two

Table 1
Summary of Survey Population, Sample, and Respondents by Sample Strata

	Population		*Sample*		*Respondents*	
Strata	*Districts*	*Schools*	*Districts*	*Schools*	*Districts*	*Schools*
Small	44	269	14	96	13	65
Medium	23	501	7	154	7	99
Large	8	400	4	156	4	108
Total	75	1,170	25	406	24	272

methods of analysis were used; first, Cronbach alphas were calculated to measure the internal consistency of the scales; and, second, a confirmatory factor analysis to test validity of the model and scales by comparing the empirical factors generated by the survey with the constructed factors.

Internal Consistency of Scales

Cronbach alphas were calculated for each of the five independent variable scales and the four outcome measures as well as the composite scales of transformation, transaction, and effects (composite of the four outcome measures). These findings are presented in Table 2. Internal consistency for all the variables ranged from 0.92 to 0.55. The lowest alpha computed was for the independent variable scale of contingent reward (0.55), which contained only three items and was the least consistent of the independent scales. Composite scales of transformation (charisma/inspiration, intellectual stimulation, individual consideration) and effects (school effects, teacher effects, program and instruction effects, student effects) had alphas >0.9. A computed alpha of 0.25 for transaction indicated that grouping together items from the contingent reward and management-by-exception scales did not form one single coherent and reliable composite scale.

Table 3 presents correlations among all the scales. Contingent reward and management-by-exception were negatively correlated (–0.30) indicating that leadership high on contingent reward was associated with low management-by-exception. Contingent reward correlated with both the transformation (0.69) and transaction (0.45) scales and therefore was not of itself a consistent contributor to the transactional scale. Management-by-exception correlated negatively with all other scales except transaction (0.73).

Confirmatory Factor Analysis

Bass's model suggests a structure consisting of two broad constructs represented by transformational and transactional leadership. Confirmatory factor analysis on the data showed the actual factor structure that emerged from a

Table 2
Internal Consistency of the Constructed Independent and Dependent Variable Scales

Scale	*No. of Items*	*Cronbach alpha*
Charisma/inspiration	13	0.92
Intellectual stimulation	3	0.85
Individual consideration	6	0.86
Contingent reward	3	0.55
Management-by-exception	3	0.77
School effects	5	0.76
Teacher effects	8	0.83
Program and instruction effects	5	0.67
Student effects	10	0.90
Composite scales		
Transformation	22	0.95
Transaction	6	0.25
Effects	28	0.93

Table 3
Intercorrelations of Scales

	Schoole	*Teache*	*Progine*	*Stude*	*Charins*	*Intstim*
Schoole	1.00	.69	.68	.71	.28	.30
Teache	.69	1.00	.74	.82	.20	.20
Progine	.68	.74	1.00	.69	.25	.24
Stude	.71	.82	.69	1.00	.18	.20
Charins	.28	.20	.25	.18	1.00	.80
Intstim	.30	.20	.24	.20	.80	1.00
Indcon	.27	.27	.31	.21	.82	.71
Conrew	.21	.27	.31	.27	.65	.59
Manbex	−.21	−.04	−.08	−.12	−.44	−.51
Transfor	.30	.24	.29	.21	.98	.86
Transact	−.01	.20	.20	.12	.11	−.03
Effects	.85	.91	.87	.93	.24	.25

	Indcon	*Conrew*	*Manbex*	*Transfor*	*Transact*	*Effects*
Schoole	.27	.21	−.21	.30	−.01	.85
Teache	.27	.27	−.04	.24	.15	.91
Progine	.31	.31	−.08	.29	.20	.87
Stude	.21	.27	−.12	.21	.12	.93
Charins	.82	.65	−.44	.98	.11	.24
Intstim	.71	.59	−.51	.86	−.03	.25
Indcon	1.00	.69	−.35	.91	.18	.28
Conrew	.69	1.00	−.30	.69	.45	.28
Manbex	−.35	−.30	1.00	−.44	.73	−.15
Transfor	.91	.69	−.44	1.00	.12	.27
Transact	.18	.45	.73	.12	1.00	.11
Effects	.28	.28	−.15	.27	.11	1.00

Scales are abbreviated as follows: Schoole, School effects; Teache, Teacher effects; Progine, Program and Instruction effects; Stude, Student effects; Charins, Charisma/inspiration; Intstim, Intellectual stimulation; Indcon, Individual consideration; Conrew, Contingent reward; Manbex, Management-by-exception; Transfor, Transformation; Transact, Transaction; Effects, School outcomes in total.

principal factors extraction when oblimin rotation into two factors was specified by the researcher. Principal factors extraction was chosen because it is widely used and conforms to the factor analytic model in which common variance is analyzed with unique and error variance removed. The decision to use oblique rotation was made because it allows the factors to be correlated (Carroll, 1983). Table 4 gives the factor loadings and structure for the two-factor solution as well as the correlations among the factors.

The actual factors emerging from the two-factor solution presented in Table 4 readily confirm the transformational scale as factor one, which includes most of the items used for that constructed factor. Two items (8 and 10) loaded only, but negatively, on factor two. Five items (7, 9, 14, 15, 16) loaded positively on factor one and negatively on factor two, and only one item (1) loaded weakly (< 0.3) on both factors. It is important to note that all three items defining the transactional construct of contingent reward loaded unequivocal-

Table 4
Factor Loadings,* Eigenvalues, Percents of Variance, and Correlations Between Factors for Principal Factors Extraction and Oblimin Rotation for Two Factor Solution for Leadership Variables on School Survey Items

delta = 0 for oblique rotation	F1	F2
Eigenvalues	12.74	1.92
Percent Variance	45.5	6.8
Cumulative Percent Variance	45.5	52.4
Leadership Items—2 Factor Solution		
Transformational Leadership		
Charisma/inspiration		
1. Insists on only the best performance from us		
2. Has capacity to overcome any obstacle	.75	
3. Leads by doing than simply by telling	.64	
4. Commands respect from everyone	.73	
5. Symbolizes success within our profession	.84	
6. Excites us with visions	.64	
7. High expectations for us as professionals	.37	–.44
8. Will not settle for second best in performance		.–47
9. Encourages us to be team players	.51	–.37
10. Less own needs more school objectives		–.31
11. Makes us feel/act like leaders	.82	
12. Provides good models to follow	.88	
13. Gives sense of overall purpose	.83	
Intellectual stimulation		
14. Challenges us to reexamine assumptions	.31	–.41
15. Prompts us to think about what we are doing	.54	–.41
16. Stimulates us to rethink the way we do things	.46	–.51
Individual consideration		
17. Facilitates support from external personnel	.60	
18. Compliments when we do outstanding work	.75	
19. Provides training for relevant skills	.59	
20. Treats us as individuals with unique needs	.70	
21. Thoughtful of our personal needs	.75	
22. Provides special recognition for good work	.79	
Transactional Leadership		
Contingent reward		
23. Assures us we can get what we want for effort	.37	
24. Provides relevant information	.52	
25. Helps us get what we decide we want	.71	
Management-by-exception		
26. Asks only of us what is essential to get work done		.67
27. Satisfied with performance if old ways work		.67
28. No change if going all right		.60
Correlations between factors	*F1*	*F2*
Factor 1	1.00	
Factor 2	–.57	1.00

*Only factor loadings > 0.30 are given

ly on factor one. This results in the constructed transactional scale being empirically defined by items 26, 27 and 28. Eigenvalues and percent variance indicate that together these two factors account for half the variance that exists in the items defining the factors with the first factor making the substantial contribution. The two-factor solution is supported strongly by the empirical data, but the characteristics of these two factors do not conform to Bass's hypothesized relationships.

Canonical Correlation

The results for the testing of successive latent roots in the canonical correlation analysis of the survey data comprising five predictor variables and four criterion variables are presented in Table 5. From the analysis reported in the table, it is clear that the first two canonical correlations are almost certainly significant even if allowance could be made for the fact that the schools do not form a simple random sample. Thus the predictor variables are related to the criterion measures in two significant ways.

Table 6 records the transformation weights and the structure coefficients for the two sets of variates related by significant canonical correlations. The transformation weights and the structure coefficients recorded in Table 6 also cannot be tested for statistical significance, but a conventional guide would be to consider only those structure coefficients in excess of 0.30 as being of sufficient magnitude to be worthy of consideration. Likewise, transformation weights equal and less than 0.20 are regarded as trivial and not of sufficient size to warrant consideration. Generally, transformation weights in excess of 0.10 are associated with more than 1% of variance explained.

According to Keeves (1986) although both transformation weights and structure coefficients are useful, experience in interpreting canonical factors has shown that a more meaningful view is frequently obtained from examining the structure coefficients when interpreting relationships between the original variables and the derived canonical variates. Tatsuoka (1973) makes this clear in the following comment:

> That this is a more reasonable approach in attempting to *interpret* the canonical factors (as against assessing the relative contribution of each original variable to the factors) becomes obvious when we recall that the standardized weights (canonical or discriminant) are partial coefficients with the effects of other variables removed or controlled. This is fine when the purpose is to gauge the

Table 5
Canonical Correlations Between Predictors and Criteria

No. of roots removed	Canonical R	R^2	Approx F	df	*P*
0	.33	.11	3.07	20	<.001
1	.28	.08	2.63	12	<.019
2	.18	.03	1.78	6	NS
3	.10	.01	1.22	2	NS

NS = Not Significant

Table 6
Transformation Weights and Structure Coefficients for Canonical Analysis Between Predictors and Criteria

Predictor measures	*Transformation weights*[a] $U_1(TF)$	$U_2(TA)$	*Structure Coefficients*[b] $U_1(TF)$	$U_2(TA)$
Charisma/inspiration	−.20	−.60	.80	−.37
Intellectual stimulation	.32	−.36	.82	−.43
Individual consideration	.59	.18	.93	−.11
Contingent reward	.41	.95	.87	.32
Management-by-exception	−.01	.53	−.41	.64
Variance extracted			.62	.17
Criterion measures	V_1	V_2	V_1	V_2
School effects	.35	−1.32	.83	−.49
Teacher effects	.04	.23	.04	.46
Program and instruction	.74	.65	.97	.21
Student effects	−.01	.61	.76	.17
Variance extracted			.55	.13
Canonical R			.33	.28
Canonical R^2			.11	.08
Redundancy			.06	.01

[a]Transformation weights > 0.20 are underlined
[b]Structural coefficients > 0.30 are underlined

> contribution of each variable in the company of others, but is inappropriate when we wish to give substantive interpretations. (p. 280)

The relationship between the original observed variables and the first pair of canonical variates may be seen from an examination of the first set of structure coefficients for the criterion variables. The first canonical criterion factor is related to, largely, program and instruction effects, school effects, and student effects. The second canonical criterion variate is positively concerned with teacher effects and negatively with school effects.

Table 6 also presents the proportion of variance extracted from the criterion and the predictor sets of variables by the two pairs of significant canonical variates, as well as the canonical correlations and the redundancy of the variance of the criteria, given the predictors. The first canonical variate has associated with it 55% of the variance of the criterion measures, and 6% of this criterion variance is explained by the predictor variables through this first canonical variate. The second canonical variate has associated with it only 13% of the variance of the criterion measures. Furthermore, less than 1% of this criterion variance is accounted for by the predictor variables through the second canonical variate. Although the two canonical variates extract over 79% of the variance from the predictor set, the canonical correlations are only moderate and reveal only a moderate leadership effect on outcomes. However, they represent a significant association that is meaningful for the model being tested and that provides considerable understanding of the interrelations between the two sets of measures.

To gauge the contribution of the predictor variables to the latent variate, transformation weights are used. The first predictor canonical variate is re-

lated to several leadership variables. It has a small negative weight for charisma/inspiration and larger positive weights for intellectual stimulation, individual consideration, and contingent reward. These results indicate that after controlling for the other characteristics, charisma/inspiration contributes negatively to forming the latent transformational variate (–.20). However, the structure coefficient (.80) indicates that charisma/inspiration is strongly related to the latent transformational variate once formed. In the school context, charisma/inspiration appears not to be important in contributing to the construct of transformational leadership, although, once leadership is identified as transformational it is strongly related to charisma/inspiration. The first criterion canonical variate has the largest structure coefficients for program and instruction effects and school effects and a strong loading on student effects. The analysis shows that the first pair of canonical variates involves transformational leadership. The variates are moderately correlated with each other (R_c=.33) and involve contingent reward insofar as it relates to transformational leadership.

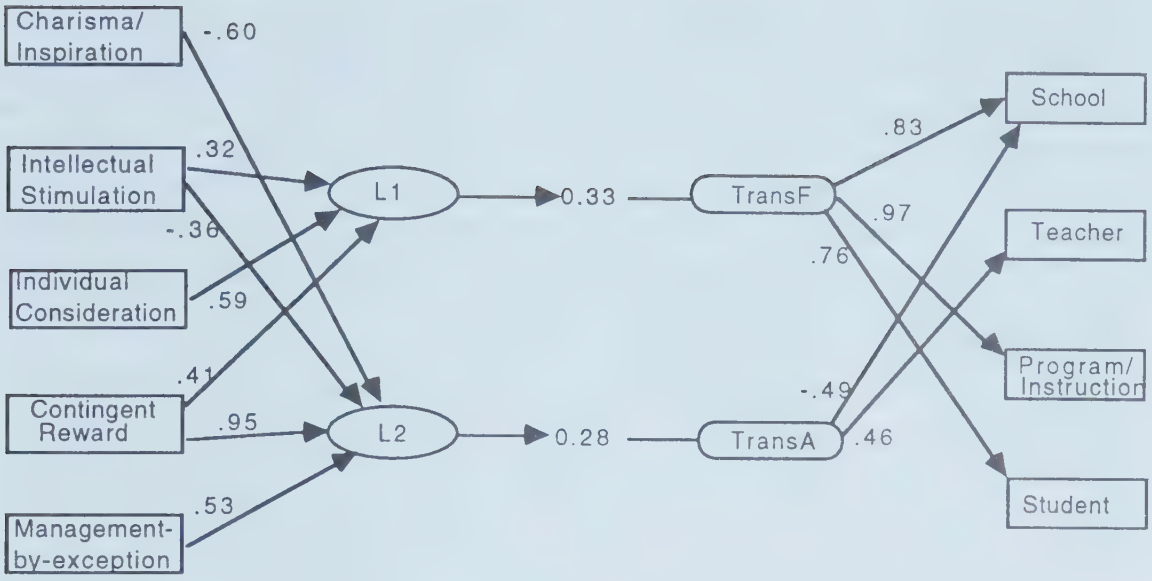
The second pair of canonical variates are also moderately correlated with each other (R_c=.28) and relate to transactional leadership. An examination of the structure coefficients shows that the criterion variate is largely and positively associated with teacher effects, while the predictor variate links strongly with management-by-exception and more moderately with contingent reward. However, in the formation of the second predictor canonical variate, after its effects on the formation of the transformation leadership variate have been partialed out, contingent reward has a large transformation weight of .95 relative to the weight of management-by-exception of .53. Two of the transformational measures (charisma/inspiration and intellectual stimulation) are also related moderately but negatively with the second predictor variate.

Figure 1 gives the path diagram for this canonical correlation analysis following an approach advanced by Van de Geer (1971) and presented in Keeves (1975). The five predictor variables give rise to two latent leadership variables L_1 and L_2 concerned with transformational and transactional leadership respectively. The path coefficients are given by the transformation weights and only path coefficients greater than 0.20 are shown in the figure. The latent variables determine in part the latent transformational variate (TransF) and the latent transactional variate (TransA) with path coefficients equal to the canonical correlations.

In the calculation of the causal paths between the latent transformational and transactional variates and the criterion measures of school outcomes, the structure coefficients are employed. Only those paths whose structure coefficients exceed 0.30 have been recorded in Figure 1. The latent transformational variate determines teacher perceptions of program and instruction effects, school effects and student effects, but is unrelated to teacher effects. The latent transactional variate influences school effects negatively and teacher effects positively. Thus the favorable influence on teacher effects is contrasted with an unfavorable influence on school effects.

Discussion

According to Bass (1985), transactional leadership relies on an exchange relationship between leader and followers, while transformational leadership sets



Only path coefficients p > 0.20 and structural coefficients f > 0.30 are presented in the figure.

Figure 1. Path diagram for canonical correlation analysis of leadership influence on school outcomes model.

aside the self-centered interests of the followers to bring about enhanced performances and change. Examination of the relationship between the dependent and independent variables using canonical regression indicates that the strongest relationship between the two sets of variables involves a transformational leadership construct defined by intellectual stimulation, individual consideration, and contingent reward. Such an outcome is contrary to a model that attempts to differentiate between transformational and transactional leadership in terms of exchange theory, insofar as contingent reward is not uniquely related to either transformational or transactional leadership. However, after the effects of contingent reward that influence transformational leadership have been removed, the residual strongly influences transactional leadership.

Factor analysis of the survey data in this study indicated that contingent reward and management-by-exception do not combine to form the construct of transactional leadership. Contingent reward correlates more strongly with each of the transformational factors than management-by-exception. Management-by-exception correlates negatively with all other variables. The operational and conceptual problems exhibited by the contingent reward scale hampers interpretation of the transactional construct and, to a lesser extent, the transformational construct.

Bass (1985; Bass & Avolio, 1988, 1990) clearly associates transactional leadership with managerial practices such as organizing, planning, and using rules and procedures, but these managerial elements are not used to operationally define transactional leadership. Although the source of the problems with the transactional scale may be a conceptual one, the low internal consistency of the composite scale indicates an operational problem, particularly when the defining factors demonstrate a satisfactory internal consistency.

Transformational and Transactional Constructs in Education

The transformational composite scale appears to provide a useful operational foundation for exploring leadership in education. There is support for the compelling argument that school leaders can promote change more successfully through employing methods associated with transformational leadership. The implication is that reliance on given procedures, rules, or reward systems is less effective. Personal dynamics, communication through ideas and example, and a perspective focusing on the group and collaboration positively influence school outcomes and facilitate school reform.

Burns (1978) has suggested that the dominant quality of transformational leadership is a developmental orientation toward followers. If school outcomes are best predicted by a developmental attitude toward followers, then a measure emphasizing individualized consideration and intellectual stimulation factors should be developed (Kirby, 1991), including the more elevating aspects of contingent reward. The correlations of individualized consideration, intellectual stimulation, and contingent reward with each of the school outcomes (Table 3) were for the most part higher than for charisma/inspiration, supporting such a refinement of this construct. Aspects of transformational leadership such as providing a model, employing collaborative problem solving, encouraging teacher development, and providing recognition and support may need to be emphasized because they appear to be more fundamental in the school setting. Nevertheless, it is evident that transformational leadership is a useful construct for educational leaders in their striving for educational reform.

In this study, transactional leadership did not prove to be as useful a component of educational leadership leading to change and improvement within schools. Although, as has been argued, this may be due to an inadequate definition of the transactional scale with an inappropriate emphasis on the exchange aspect. The results of this study support conclusions expressed by King (1989) with respect to the transactional construct. She suggested that the exchange relationship defining transactional leadership may become a component of transformational leadership when used in a

> charismatic or inspirational manner. The intricacies of the individual exchange may determine whether or not it indicates transformational leadership. Extra effort in exchange for a promotion may characterize transactional leadership. However, extra effort in exchange for personal growth, increased self-esteem, or identification with a cause greater than oneself indicates transformational leadership. (pp. 160-161)

Additional insight into the possible nature of a transactional scale can be derived from the canonical regression analysis and the strong but negative correlations of management-by-exception with the transformational scales. Management-by-exception refers to a more passive style of leadership, one that may be seen as giving over more control to followers. The strong association of this variable with teacher effects implies that "less involved" leadership may enhance teacher outcomes although it may have a negative impact on school effects. Transactional leadership in education may include an aspect of leader uninvolvement that may result in contradictory outcomes: empowerment of teachers at the expense of enhanced outcomes for the whole organiza-

tion. Such empowerment would not occur because of active leader behavior, but from leadership withdrawal, perhaps enacted in a climate of frustration and abandonment. As with the exchange relationship, the effects of leadership involvement may be context specific. It is not the presence or absence of involvement that defines leadership, but the nature of the involvement that determines whether it is transformational or transactional.

Usefulness of the Model for Explaining School Improvement Outcomes

Two latent variables were derived from canonical analysis of the survey data in this study. One was identified as transformational leadership and was primarily characterized by individual consideration, contingent reward, and intellectual stimulation. Charisma/inspiration was a marginal and negative contributor to this construct. This implies that the characteristics of leadership represented by charisma/inspiration do not discriminate between leaders within the school setting. Perhaps in the setting under survey too little variation exists in school leaders between their charisma/inspiration characteristics, or strong charismatic characteristics may not be present in sufficient numbers of school leaders, as has been implied (King, 1989; Kirby, 1991). Leadership characteristics defined by contingent reward related both to the transformational and transactional constructs. This implies that some aspects of the exchange relationship are of an elevated nature leading to higher levels of performance motivated by goals for the greater good. However, the nature of contingent reward is likely to be context specific and take on transactional characteristics at other times. The second latent variable derived from canonical analysis was identified as transactional leadership and was characterized by contingent reward and management-by-exception, as well as low measures on charisma/inspiration and intellectual stimulation.

These empirically derived constructs had the following effects: transformational leadership influenced all school outcomes except teacher effects; transactional leadership influenced teacher effects favorably and school effects unfavorably. It can be argued that in the school context, transformational leadership if defined by intellectual stimulation, individual consideration, and contingent reward influences a broad range of school outcomes that does *not* include teacher effects. The relationship of the transformational leadership construct to the charisma/inspiration variable is less clear. Although the variable is correlated with the construct, it does not contribute positively to its formation. Teacher effects are influenced by transactional leadership when it is defined as high on contingent reward and management-by-exception and low on intellectual stimulation and charisma/inspiration, as might be expected. The implication is that leadership impacts on teachers when school leaders are less involved, less concerned with goal setting and planning, and more exchange theory oriented. At the same time, however, such school leaders impact on the whole school negatively. When transactional leadership characteristics are more dominant, then teacher effects are enhanced and perception of school effects are depressed. To enhance teacher effects requires less involved and more passive leaders. Possibly such leaders allow teachers to take more responsibility (deliberately or by default) and experience a greater sense of control over their environment (although it does not lead to positively perceived school effects). Perhaps teacher power and control leads to en-

hanced teacher effects, whereas leader power and control enhances school, program and instruction, and student effects. It follows that shared power must broaden the impact of leadership on school outcomes. Perhaps transactional leadership conveys the idea of more teacher power, whereas transformational leadership conveys the idea of more leader power. In the current climate of teacher empowerment for school reform, these leadership constructs and these analyses indicate the problems involved.

Another important outcome of canonical analysis relates to the measure of redundancy in the criteria accounted for by the predictors. Only 6% of the variance in outcomes is explained by the predictors in the first pair of canonical variates, which represents the strongest (but moderate) association extracted from the data, and only a further 1% by the second pair. A total of only 7% of the variance in school outcomes is explained by the leadership variables. Although there is empirical support for meaning in the factors used and the way they are operating, there is a lot of unexplained variance in the two sets of variables. In the school setting, leadership defined in terms of transformational and transactional constructs is only a weak predictor of the educational outcomes examined. Data used in this study indicate that there may be more important constructs operating on educational outcomes than leadership as construed in Bass's model. It is possible that a redefinition of transformational and transactional leadership may increase the explanatory power of these concepts. Alternatively, more accurate measurement of the outcomes may lead to better explanations.

Recent research by Leithwood and Jantzi (1990) and Leithwood, Jantzi, Silins and Dart (1992) indicate that school leadership is mediated in its effects on school and student outcomes by in-school processes such as school goals, school culture, and teachers. Direct impact of leadership on school and student outcomes was found to be insignificant, whereas the mediated effect was quite significant. In these studies, a broader definition of transformational leadership (Podsakoff, Mackenzie, Moorman, & Fetter, 1990) was used than the one operationalized by Bass. However, there is evidence from this study to suggest that educational leadership has a significant indirect influence on school outcomes in the context of reform. More needs to be known about the mediators of leadership in general and transformational leadership in particular.

Conclusion

This study tested Bass's (1985; Bass & Avolio, 1988) model of transformational and transactional leadership applied to the school setting and found it provides support, in general, for a two-construct model. However, there is considerable overlap between the factors defining the two constructs so that their usefulness in explaining differences in school improvement outcomes is somewhat confounded. The transformational construct holds the most promise as a predictor of enhanced school outcomes. This construct appears to be a compendium of personal qualities of the leader that are not captured well by the charisma/inspiration factor, and leader behaviors represented by the factors of intellectual stimulation and individual consideration, as well as contingent reward. The transactional construct could be more useful in explaining leadership in education if redefined in terms of managerial and instructional tasks, together with teacher empowerment rather than exchange

theory. This is in line with the general discussion of differences between managers and leaders that has gone on for some time (Zaleznik, 1977) and the current teacher empowerment movement (Joyce, 1990).

Further research is needed to understand the behaviors that result in a leader being perceived as transformational in the school context. Operationalizing this construct in education more clearly will provide guidelines for leadership training and for orchestrating school restructuring. Leadership research in the school setting needs to investigate the implied contradictory nature of improving teacher and school effects. It seems likely that teachers should be important mediators of the impact of leadership in any school, so that models of leadership postulating direct influences on school outcomes are not as useful in educational settings. Regardless of the leadership constructs employed in studying school improvement, a more fundamental problem may exist if the model of leadership for school reform does not take into account processes within schools that operate to mediate the impact of school leadership on school outcomes.

Multivariate analysis of survey data has indicated that school leadership as presently construed may have little explanatory power for school outcomes. Bass's model assumes a direct effect of leadership on outcomes. It has been suggested that leadership operates indirectly to enhance school outcomes and that the mediated effect is stronger than the direct effect. Refining the transformational construct and increasing understanding of the indirect as well as the direct effects of leadership should help to explain the differences in outcomes of school improvement programs.

References

Avolio, B., & Bass, B. (1987). Transformational leadership, charisma, and beyond. In J. Hunt, B. Baliga, H.P., Dachler, & C. Schriesheim (Eds.), *Emerging leadership vistas* (pp. 29-50). Lexington, MA: Lexington Books.

Bass, B.M. (1985). *Leadership and performance beyond expectations*. New York: Free Press.

Bass, B.M., & Avolio, B.J. (1988). *Manual for the Multifactor Leadership Questionnaire*. Binghamton, NY: Center for Leadership Studies, State University of New York at Binghamton.

Bass, B.M., & Avolio, B.J. (1990). *Transformational leadership development: Manual for the Multifactor Leadership Questionnaire*. Palo Alto, CA: Consulting Psychologists' Press.

Bass, B.M., Waldman, D.A., Avolio, B.J., & Bebb, M. (1987). Transformational leadership and the falling dominoes effect. *Group and Organization Studies, 12*(1), 73-87.

Beare, H., Caldwell B.J., & Millikan, R.H. (1989). *Creating an excellent school: Some new management techniques*. New York: Routledge.

Burns, J.M. (1978). *Leadership*. New York: Harper & Row.

Carroll, J.B. (1983). Studying individual differences in cognitive abilities: through and beyond factor analysis. In R.F. Dillon & R.R. Schmeck (Eds.), *Individual differences in cognition*. New York: Academic Press.

Fullan, M.G. (1991). *The new meaning of educational change* (2nd ed.). New York: Teachers College Press.

Hoover, N.R. (1988). Transformational and transactional leadership: A test of the model (Doctoral dissertation, University of Louisville, 1987). *Dissertation Abstracts International, 48*, 3020A.

Hoover, N.R., Petrosko, J.M., & Schulz, R.R. (1991). *Transformational and transactional leadership: An empirical test of a theory*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Joyce, B. (Ed.). (1990). *Changing school culture through staff development*. The 1990 Yearbook of the Association for Supervision and Curriculum Development. Alexandria: ASCD.

Keeves, J.P. (1975). The home, the school and achievement in mathematics and science. *Science Education, 59*, 439-460.

Keeves, J.P. (1986). Canonical correlation analysis. *International Journal of Educational Research, 10*, 1164-173.

King, M.I. (1989). Extraordinary leadership in education: Transformational and transactional leadership as predictors of effectiveness, satisfaction, and organizational climate in K-12 and higher education (Doctoral dissertation, University of New Orleans). *Dissertation Abstracts International, 50*, 08A.

Kirby, P.C. (1991). *Extraordinary leaders in education: A test of Burns's model of transformational leadership.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Leithwood, K.A., Begley, P.T., & Cousins J.B. (1991). The nature, causes and consequences of principals' practices: An agenda for future research. *Journal of Educational Administration, 28*(4), 5-31.

Leithwood, K., & Jantzi, D. (1990). Transformational leadership: How principals can help reform school cultures. *School Effectiveness and School Improvement, 1*, 249-280.

Leithwood, K., Jantzi, D., Silins, H., & Dart, B. (1992). *Transformational leadership and school restructuring.* Paper presented at the International Congress for School Effectiveness and Improvement, Victoria, BC.

Podsakoff, P.M., Mackenzie, S.B., Moorman, R.H., & Fetter, R. (1990). Transformational leadership behaviors and their effects on followers' trust in leader, satisfaction and organizational citizenship behaviors. *Leadership Quarterly, 1*(2), 107-142.

Royal Commission on Education. (1987-1988). *A legacy for learners: The report of the Royal Commission on Education, 1988.* Victoria, BC: Author.

Sashkin, M. (1988). The visionary principal: School leadership for the next century. *Education and Urban Society, 20*, 239-249.

Tatsuoka, M.M. (1973). Multivariate analysis in educational research. In F.N. Kerlinger (Ed.), *Review of Research in Education I* (pp. 273-319). Itasca, IL: Peacock.

Van de Geer, J.P. (1971). *Introduction to multivariate analysis for the social sciences.* San Francisco, CA: Freeman.

Waldman, D.A., Bass, B.M., & Einstein, W.O. (1987). Leadership and outcomes of performance appraisal processes. *Journal of Occupational Psychology, 60*, 177-186.

Yammarino, F.J., & Bass, B.M. (1990). Long-term forecasting of transformational leadership and its effects among naval officers. In K.E. Clark & M.B. Clark (Eds.), *Measures of leadership* (pp. 151-170). West Orange, NJ: Leadership Library of America.

Zaleznik, A. (1977). Managers and leaders: Are they different? *Harvard Business Review, 55*, 67-68.

The Alberta Journal of Educational Research *Vol. XXXVIII, No. 4, December 1992, 335-354*

Tichatonga J. Nhundu
University of Zimbabwe

Job Performance, Role Clarity, and Satisfaction Among Teacher Interns in the Edmonton Public School System

This article reports on the findings of a study that investigated the nature of the relationship between various indices of job satisfaction, role clarity, and job performance using the perceptions of a sample of teacher interns and their supervising teachers. The perceptions of teacher interns concerning job satisfaction and role clarity were determined using the Minnesota Satisfaction Questionnaire (MSQ) and a role clarity instrument developed along the lines of the Rizzo, House, and Lirtzman (1970) Role Ambiguity questionnaire, respectively. Job performance was determined from composite evaluations by dyads of supervisors and teacher interns. Generally, there was a high level of agreement between supervising teachers and interns concerning areas of lesser and greater performance. However, self ratings possessed less leniency than corresponding ratings by supervising teachers. Second, self appraisals showed significantly stronger correlations with both role clarity and job satisfaction than supervisor appraisals. Although perceptions of high levels of role clarity and job performance were reported in this study, the surveyed interns still reported experiencing lower levels of job satisfaction than the MSQ normative data for a comparable group. Finally, the findings of this study also provided tentative support for Herzberg's (Herzberg, Mausuer, & Snyderman, 1959) Hygiene theory.

Cet article présente les résultats d'une étude qui examinait la nature de la relation entre différents indices de la satisfaction au travail, de la clarté des rôles et de la performance au travail d'après les perceptions d'un nombre d'enseignant(e)s qui participaient à un programme d'internats et de leurs enseignant(e)s coopérant(e)s. Les perceptions des enseignant(e)s qui participaient à un programme d'internats concernant la satisfaction au travail et de la clarté des rôles ont été déterminé en utilisant l'instrument Minnesota Satisfaction Questionnaire (MSQ) et un instrument de clarté des rôles développé selon un questionnaire proposé par Rizzo, House et Lirtzman (1970). La performance au travail a été déterminée selon une évaluation composée par des superviseurs et des enseignant(e)s internes en relation dyadique. En général, il y avait un niveau élevé de la corrélation entre les perceptions des enseignant(e) superviseurs et internes dans les domaines de la performance. Cependant, l'autoévaluation présentait moins de satisfaction que l'évaluation correspondante par les enseignant(e)s superviseurs. Deuxièmement, l'évaluation de soi démontrait une corrélation plus significative entre les deux, la clarté des rôles et la satisfaction au travail, qu'avec les évaluations des superviseurs. Même s'il semblerait y avoir une autre corrélation entre la clarté des rôles et la performance au travail, un sondage effectué parmi les enseignant(e)s internes indique qu'il existe quand même un niveau inférieur de satisfaction au travail qu'indiqué sur les données normatives MSQ pour un groupe comparable. Enfin, les résultats de cette étude appuient tentativement la Théorie d'Hygiène de Herzberg (Herzberg, Mausuer, & Snyderman, 1959).

Introduction

This article reports on the findings of a study concerning the perceptions of teacher interns on three dimensions of their job, namely, job satisfaction, role

Tichatonga J. Nhundu is a lecturer in Educational Administration.

clarity, and job performance. Specifically, the main purposes of this study were: (a) to determine whether teacher interns experienced role clarity and to identify the variables that contributed positively and negatively to perceptions of role clarity; (b) to assess levels of job performance of teacher interns using perceptions of both teacher interns and their supervising teachers; (c) to determine whether teacher interns experienced job satisfaction and to identify factors that influenced perceptions of job satisfaction; and (d) to explore and determine the nature of the relationships linking role clarity, job performance, and job satisfaction.

The Nature of the Alberta Initiation to Teaching (Internship) Project (ITP)

On April 22, 1985 the province of Alberta, Canada, announced the establishment of the Initiation to Teaching (Internship) Project (ITP), which was introduced on an initial two year experimental basis commencing September 1, 1985 and ending June 30, 1987. According to the ITP Guidelines (Alberta Education, 1985, p. 10) the main purpose of the project was "to provide for the continued professional training of graduates of the Faculties of Education in an environment that will facilitate the transition from student to professional teacher, as well as to provide employment opportunities for those graduates who may otherwise be unemployed." Only unemployed graduates were eligible for the ITP provided they had not graduated prior to 1983. The graduates were required to serve as interns for a period of one year during which they taught in cooperating schools under the close supervision of experienced teachers. During this period the interns received about half the salary of a beginning teacher, but were not assured of employment following the completion of their internship. In addition, the period served as an intern would not be considered for tenure, salary, or grading purposes.

Research on Role Ambiguity, Job Satisfaction, and Job Performance

Role ambiguity (obverse of role clarity) generally arises from discrepancy between job-related information that is made available to the worker and the ideal information required to assist the worker to adequately perform his/her role (Farber, 1983; Kahn, Wolfe, Quinn, & Snoek, 1964; and Rizzo, House, & Lirtzman, 1970). The lack or absence of clarity concerning job-related information was attributed by Kahn et al. (1964) to organizational complexity, rate of organizational change, and managerial philosophy.

Job satisfaction, which is defined by Locke (1976, p. 1342) as a "pleasurable emotional state resulting from the perception of one's job as fulfilling or allowing the fulfillment of one's important job values," is generally viewed as an outcome of performance and not a determinant of employee productivity. Several studies, including those by Friesen, Holdaway, and Rice (1983); Holdaway (1978); Miskel (1982); and Wickstrom (1973), have focused on the determinants of job satisfaction and their implications for career development of educational personnel. Thus recognition, responsibility, sense of achievement, the work itself, and good interpersonal relationships with students were found to improve sources of satisfaction for teachers, while rapport with administrators was strongly associated with overall career satisfaction. But for the majority of teachers, Lortie (1969) and Sergiovanni (1967) found that promotion opportunities did not motivate or provide job satisfaction, while

Holdaway (1978) and Schmidt (1976) found achievement, recognition, and advancement (promotion) as the main source of job satisfaction among those teachers who aspired to be administrators. The main source of job satisfaction for school administrators included achievement, responsibility, and autonomy (Friesen, Holdaway, & Rice, 1983) while professional satisfaction for beginning teachers was reported by Hewitson (1975) to be strongly associated with higher levels of supervision and decisional involvement.

Job satisfaction is important in education as it is closely associated with teacher turnover and absenteeism. School systems stand to lose substantially in terms of productivity, low morale, and financially by having dissatisfied teachers in their staff because these teachers are less productive, less positive, and are often not available to provide continuity of instruction.

The third research variable, job performance, is seen by Van de Ven and Ferry (1980, p. 82) as a value judgment that reflects the criteria used to judge how well people or organizations achieve their operating objectives. French, Kass, and Rosenzweig (1985) identified three basic variables that affect performance as ability, effort, and opportunity. These variables are in turn dependent on several factors. For example, effort is a function of needs, goals, expectations, and rewards. Thus, if opportunity and ability are held constant, then performance will depend on effort in that lack of clarity concerning, say, needs, expectations, and goals may lead to low performance.

Several studies in educational psychology generally concur that role ambiguity is inversely related to job satisfaction. For example, Kahn et al. (1964) reported that "role ambiguity leads to increased emotional tension and to decreased satisfaction with one's job" (p. 85). Subsequent studies by Beehr (1976), Greene (1972), Greene and Organ (1973), and Keller (1975) provided further support to the findings of Kahn et al. Similarly, House and Rizzo (1972), McLean (1979), Miles (1976) and Rizzo et al. (1970) reported a dissatisfaction. However, research findings on the relationship between role ambiguity and job satisfaction have not always been supportive or conclusive. For example, studies by Hamner and Tosi (1974) and Tosi (1971) found no significant relationship between role ambiguity and job satisfaction, while Szilagyi (1977) reported that role ambiguity was "not always negatively related to job satisfaction" (p. 376). Accordingly, it was expected that the present study would help further explain and clarify the nature of this relationship.

Role ambiguity has also been reported to be inversely related to job performance. Studies by Smith (1957), Schwab and Iwanicki (1982), Kahn et al. (1964), Cherniss (1980) and Burke and Belcourt (1974), for example, have reported findings indicating that ambiguous conditions and expectations lead to significant reduction in productivity or effectiveness, directly or indirectly, due to burnout or stress, while studies by House and Rizzo (1972), Miles (1976), and later Comadema (1984) concur that role ambiguity is negatively related to job performance. Evidence has also been reported by Beehr, Walsh, and Taber (1976) linking role ambiguity to reduced levels of employee participation and job involvement. However, a thorough analysis of literature reviews carried out by Lawler and Porter (1967) concluded that job performance was also directly related to job satisfaction.

Conceptual Framework

The conceptual framework proposed for this study constitutes a working model only and was used as a basis for defining the research problem and subproblems and the development of the research methodology including data collection procedures. It was developed from a synthesis of the various relationships between the major research variables reported in the literature. The framework consists of three major research variables: job satisfaction, role ambiguity, and job performance.

A search of existing literature on the nature of the relationship between job satisfaction and performance reveals three general conceptualizations explaining this relationship (Locke, 1970; Schwab & Cummings, 1970). The first formulation is that job satisfaction leads to job performance, while the second, which generally concurs with the former, further asserts that this relationship is moderated by intervening variables. The third and more widely accepted formulation stipulates that job satisfaction is both a result of, and dependent on, good performance and not vice versa (Lawler & Porter, 1967; Pigge & Lovett 1985; and Siegel & Bowen, 1971). This conceptualization of the performance-satisfaction relationship constitutes one dimension of the proposed conceptual framework of this study.

The second dimension links role ambiguity directly to job satisfaction indicating that role ambiguity is an intrinsic cause of job satisfaction as reported by Comadema (1984), House and Rizzo (1972), Miles (1976), and Rizzo et al. (1970). However, Cohen (1959), Kahn et al. (1964), and Smith (1957) also reported that role ambiguity is directly related to job performance. The link between role ambiguity and job performance, therefore, constitutes the third and final dimension of the proposed conceptual framework for this study.

The conceptual frame may help explain why, for example, insufficient information (that is, role ambiguity) about a focal position may lead to poor job performance or low productivity, which in turn may lead to job dissatisfaction. The proposed conceptual framework also further helps to explain why a person experiencing role ambiguity may be intrinsically dissatisfied with the job. To this end, role ambiguity is a predictor variable for both job performance and job satisfaction. Similarly, job performance, which is also an intrinsic cause of job satisfaction, is a predictor variable of job satisfaction.

Sample

Fifty interns and their supervising teachers were randomly selected from a population of 150 interns employed in the Edmonton Public School system. All the respondents were contacted by telephone regarding their possible participation in the study. Participation was in dyadic pairs so that if one member of a pair declined to participate in the study, the other member of that pair would automatically be dropped. Seven interns and/or supervising teachers declined, while three interns later quit the internship program before data collection was completed. Accordingly, the final research sample was made up of 54 elementary and 26 high school teacher interns and their corresponding supervising teachers. All the respondents were asked to sign a release form assuring them of the anonymity, confidentiality, and the non-

evaluative nature of the research as well as the right to withdraw from the study at any point during the investigation.

Instrumentation

Three questionnaires were used to elicit respondents' perceptions concerning the three major research variables: role clarity, job performance, and job satisfaction. Teacher interns were required to answer all questionnaires, whereas supervising teachers answered only a 30-item job performance questionnaire identical to the one administered to teacher interns.

The questionnaire on role clarity (Appendix A) consisted of 16 items. This instrument was developed by the researcher using literature on role ambiguity, in particular the Rizzo et al. (1970) Role Questionnaire. All the items were structured on a six-point Likert scale and the respondents were requested to indicate the extent to which certain conditions existed in their job using the six-point scale of Always (6), Frequently (5), Sometimes (4), Occasionally (3), Rarely (2), and Never (1). Typical examples of items on the questionnaire were "I am certain about my responsibilities," "I know exactly how much authority I have," and "I receive feedback on how I do my work."

Items on the job performance questionnaire (Appendix B) administered to teacher interns and supervising teachers were identically worded. The 30 items on this questionnaire were structured on a five-point (1-5) Likert-like scale by the researcher after consulting several instruments used in the Edmonton Public School system and the University of Alberta's Faculty of Education to evaluate teacher interns and student teachers respectively. In addition, literature on teacher effectiveness and evaluation was consulted and greatly assisted in constructing the questionnaire. The 5-point performance rating scale used was calibrated as follows: 1 (poor), 2 (fair), 3 (good), 4 (very good), and 5 (exceptional). Means and standard deviations were also computed for responses to this questionnaire.

Both job performance and role clarity questionnaires were pilot tested using a sample of former teacher interns who had come out of the same program the previous year. Reliability coefficients of 0.89 and 0.81 for role clarity and job performance were obtained using split-half correlational analysis. Validity of the instruments was based on face validity with the concurrence of several experienced researchers.

The Minnesota Satisfaction Questionnaire (Long Form) developed by Weiss, Dawis, England, and Lofquist (1967) was used to collect data on job satisfaction of teacher interns. Permission was sought and granted by the Vocational Psychology Research Center of the University of Minnesota, and 50 copies of the instrument were purchased for use in the study.

Procedure

The major reason for using the questionnaire approach in this study was the anonymity afforded to respondents. According to Mouly (1981), this approach allows the respondents to express their views anonymously, views that may otherwise be opposed or influenced by others or by researcher bias. This was an important consideration in this study because the researcher did not want the interns' ratings of their performance to be influenced by their supervisors. However, a semistructured interview schedule was also used to complement

the questionnaire and provide multiple perspectives and also to allow for deeper understanding of the variables under investigation.

All questionnaires were hand delivered in order to boost the rate of return and also to give the researcher the opportunity to establish rapport with respondents and further explain the purpose of the study. The respondents were asked to complete the questionnaires and mail them separately to the researcher using stamped and self-addressed envelopes provided by the researcher.

Research Findings

Question 1 : Extent of Role Clarity Experienced by Teacher Interns

For each of the 16 tasks, means and standard deviations were computed based on the responses of teacher interns. Table 1 contains a list of the 16 tasks ranked according to levels of perceived clarity. In addition, Table 1 also shows the proportion of interns who responded in each of the four categories on the rating scale. Of the 16 tasks, six were classified Role Expectations (RE), eight as Role Responsibilities (RR), and two as Performance Feedback (PF).

The profile of responses reflected in Table 1 clearly shows that there was general consensus among the respondents indicating that they experienced role clarity concerning their job. The overall modal response for the 16 items was "always." Sixty-six percent of the interview sample reported that they found their roles always clear compared with 17% who said their roles were "frequently" clear, 9% who reported that their roles were "sometimes" clear, and 8% who "rarely" found their roles clear.

An examination of Table 1 shows that of the eight highest rated tasks according to level of perceived clarity, two were in the area of Role Expectations while six were classified as Role Responsibilities. Of the bottom eight tasks, two were Role Responsibilities, two were Performance Feedback, and four were Role Expectations. This shows that the areas that contributed most to role clarity were those concerned with Role Responsibilities of teacher interns, which constituted six out of the eight highest rated clarity items followed by items on Role Expectations, while those on Performance Feedback contributed least to role clarity.

Interview data that generally corroborated questionnaire data provided more insight into the perceptions of respondents concerning performance feedback. The interviewed interns expressed concern and anxiety over lack of supervisory feedback. The general feeling among these interviewees was that although they were supervised or observed in class, they were otherwise not provided with feedback following these class visits. For example, one female elementary intern who summed up the general feeling among her colleagues complained that although her supervisor often sat in her class, "She doesn't tell me at the end of the lesson how I was doing. I don't even know for sure how she feels about my teaching."

The prominence of feedback as a possible source of role ambiguity highlighted in this study is supported by studies by Kahn et al. (1964) and Rizzo et al. (1970) who first identified uncertainty about the consequences or responses of others to one's behavior as an indication of role ambiguity. Subsequent studies by Pigge and Lovett (1985) and Walsh, Taber, and Beehr (1980) further revealed that performance feedback was not only a source of role ambiguity,

Table 1
Perceptions of Teacher Interns Ranked According to Mean Ratings of Extent to which they Experienced Role Clarity on 16 Dimensions (N=40)

Rank	*Category**	*Mean*	*Tasks*	*Rarely*	*Sometimes*	*Frequently*	*Always*
1	RE	3.725	I know exactly where to go for help				
			Frequency	2	2	1	35
			Percentage	5.0	5.0	2.5	87.5
2	RR	3.650	I have clear goals and objectives for my work				
			Frequency	2	2	4	32
			Percentage	5.0	5.0	10.0	80.0
3	RR	3.625	I am certain about my responsibilities				
			Frequency	1	4	4	31
			Percentage	2.5	10.0	10.0	77.5
4	RR	3.600	Terms and conditions of service are clear				
			Frequency	2	4	2	32
			Percentage	5.0	10.0	5.0	80.0
5	RR	3.575	My rights and privileges are clear				
			Frequency	1	4	6	29
			Percentage	2.5	10.0	15.0	72.5
6	RR	3.550	I know how I am expected to work with colleagues				
			Frequency	3	3	3	31
			Percentage	7.5	7.5	7.5	77.5
7	RR	3.525	I know exactly how much authority I have				
			Frequency	2	5	3	30
			Percentage	5.0	12.5	7.5	75.0
8	RR	3.475	Clarity of school jurisdiction policies and regulations				
			Frequency	4	2	5	29
			Percentage	10.0	5.0	12.5	72.5
9	RE	3.425	I get clear and explicit directions and orders				
			Frequency	4	2	7	27
			Percentage	10.0	5.0	17.5	67.5
10	RR	3.400	I know who to report to				
			Frequency	2	5	8	25
			Percentage	5.0	12.5	20.0	62.5
11	RR	3.350	I do not ask others about what I am supposed to do				
			Frequency	2	5	10	23
			Percentage	5.0	12.5	25.0	57.5
12	RE	3.325	I get clear explanations of what to do				
			Frequency	1	6	12	21
			Percentage	2.5	15.0	30.0	52.5
13	RE	3.175	I am certain about the acceptability of my work				
			Frequency	5	2	14	19
			Percentage	12.5	5.0	35.0	47.5
14	RE	3.150	I am certain about how I will be evaluated				
			Frequency	5	3	13	19
			Percentage	12.5	7.5	32.5	47.5
15	PF	3.125	I receive feedback when I expect it				
			Frequency	7	2	10	21
			Percentage	17.5	5.0	25.0	52.5

Table 1 (continued)

Rank	Category*	Mean	Tasks	Rarely	Sometimes	Frequently	Always
16	PF	2.850	I receive feedback on how to do my work				
			Frequency	6	8	8	18
			Percentage	15.0	20.0	20.0	45.0
Overall Frequency				49	59	110	422
Overall Mean				3.063	3.688	6.875	16.475
Overall Percentage				7.656	9.219	17.188	65.938

* Categories:
RE = Role Expectations
RR = Role Responsibilities
PF = Performance Feedback

but that it was also an important predictor of job satisfaction. Thus although the research sample in this study reported generally high levels of role clarity, great concern and anxiety generated by lack of feedback is an important finding that highlights the importance of performance feedback for both job performance and job satisfaction.

Self And Supervisor Assessed Performance

Question 2. Perceptions of Teacher Interns on the Effectiveness of their Performance. Self appraisals reported here are those job performance ratings that reflected the perceptions of teacher interns on how well they did on selected dimensions of their job, while supervisor appraisals are similar to ratings reflecting the perceptions of supervising teachers concerning the performance of the teacher interns on the same job dimensions. Means and standard deviations were computed for each of the 30 tasks based on the responses of teacher interns and their supervising teachers using a five-point rating scale ranging from 1 (poor), 2 (fair), 3 (good), 4 (very good), to 5 (exceptional). An overall mean score of 3.85 obtained indicates that on a five-point scale, teacher interns perceived their performance as a little more than just "good," while supervising teachers rated the performance of interns at a mean score of 4.00, that is, "very good" on the same five-point scale.

Table 2 lists the top 10 tasks ranked according to perception of teacher interns concerning the extent to which they performed these tasks. Of the 10 highest rated tasks, eight were classified as "Human Relations" (HR), one as "Curriculum and Instruction" (CI), and one as "Personal Development" (PD).

Table 3 contains a list of the bottom 10 tasks according to the interns' perceptions of the extent of their performance and shows that five areas were classified as "Curriculum and Instruction" (CI), three as "School-Community Relations" (SCR), and one each as "Personal Development" (PD) and "Human Relations" (HR).

It is evident from Table 2 that the 10 highest rated job performance tasks suggest that the greatest concern for the interns was establishing and maintaining good human relations. This high preference for human relations or ideographic dimensions of their job may well reflect the interns' feelings of job security that influenced them to build positive relations with other members

of the school with the hope of improving their chances of being asked back the following year. This analysis is supported by interview data where the interview sample showed great concern for job security, especially because they were not assured of employment following the completion of the internship period. As one junior high school intern aptly put it: "I have to impress them, hoping to be asked back next year. It is therefore important to me to build good relations with them, some form of PR."

Table 2
Top Ten Tasks Ranked According to Teacher Interns' Ratings of How Well They Thought They Were Performing Their Job (N=40)

Rank	*Category**	*Task*	*Mean*
1	HR	Maintains good rapport with supervisors	4.28
2	HR	Shows empathy for students	4.25
3	HR	Maintains good rapport with colleagues	4.23
4	HR	Cooperates with colleagues	4.15
5	HR	Good leadership qualities	4.13
6	CI	Reinforcing students	4.10
7	HR	Student supervision	4.05
8	HR	Reflects and acts on supervisory advice	4.04
9	PD	Ability to make independent decisions	4.03
10	HR	Ability to make friendship with colleagues	3.98

*Categories:
HR = Human Relations
CI = Curriculum and Instruction
PD = Personal Development

Table 3
Bottom Ten Tasks Ranked According to Teacher Interns' Ratings of How Well They Thought They Were Performing Their Job (N=40)

Rank	*Category**	*Task*	*Mean*
21	CI	Class management and control	3.63
22	SCR	Involves parents in learning of kids	3.61
23	CI	Deals with individual differences and needs of students	3.60
24	CI	Appropriateness of lesson introduction and closure	3.58
25	CI	Preparation of long and short term plans	3.53
26	PD	Student counseling	3.49
27	CI	Appropriateness of questioning techniques	3.40
28	HR	Participation in staff meetings	3.18
29	SCR	Holding parental conferences	3.12
30	SCR	Participation in community activities	2.76

*Categories:
CI = Curriculum and Instruction
SCR = School-Community Relations
PD = Personal Development

The teacher interns' concern for the human relations dimension of their job is further revealed in Table 3 where only one of the 10 human relations-related tasks appears among the bottom rated 10 job performance tasks. Interestingly, Table 2 also shows that there were no tasks from the area of school community relations among the top 10 tasks ranked highest by teacher interns. All three school-community relations tasks were ranked among the bottom 10 tasks, with two of the tasks receiving the least performance effectiveness rating by teacher interns.

Question 3: Perceptions of Supervising Teachers on the Effectiveness of the Interns' Performance.

The rating of supervising teachers concerning how well teacher interns were performing was rank ordered to determine areas of greater and lesser performance. The top 10 tasks that received highest rating scores by supervising teachers appear in Table 4. According to Table 4, eight of the 10 highest rated tasks were classified as Human Relations (HR), one as Curriculum and Instruction (CI), and the other as Personal Development (PD). No tasks in the area of School-Community Relations (SCR) appeared among the top 10 tasks.

The bottom 10 tasks ranked according to the supervising teachers' perceptions of the performance effectiveness of teacher interns are listed in Table 5. Of the bottom 10 tasks, five were classified as Curriculum and Instruction (CI), three as School-Community Relations (SCR), and one each as Personal Development (PD) and Human Relations (HR). As in the case of the teacher interns' bottom 10 tasks, only one task in the area of Human Relations appeared on the list.

Table 4
Top Ten Tasks According to Supervising Teachers' Ratings of How Well They Thought Teachers Interns Were Performing Their Job (N=40)

Rank	*Category**	*Task*	*Mean*
1	HR	Reflects and acts on supervisory advice	4.38
2	HR	Maintains good rapport with supervisors	4.33
3	HR	Maintains rapport with colleagues	4.31
4	HR	Shows empathy for students	4.25
5	HR	Good leadership qualities	4.23
6	HR	Consistence and fairness with students	4.22
7	HR	Cooperates with colleagues	4.21
8	HR	Ability to make friendship with colleagues	4.20
9	CI	Suitability of student activities	4.13
10	PD	Shows initiative	4.08

*Categories:
HR = Human Relations
CI = Curriculum and Instruction
PD = Personal Development

Table 5
Bottom Ten Tasks Ranked According to Supervising Teachers' Ratings of How Well They Thought Teacher Interns Were Performing Their Job (N=40)

Rank	*Category**	*Task*	*Mean*
21	CI	Appropriateness of objectives	3.85
22	CI	Deals with individual differences and needs of students	3.80
23	CI	Appropriateness of lesson introduction closure	3.78
24	PD	Student counseling	3.77
25	SCR	Participation in community activities	3.74
26	CI	Class management and control	3.70
27	CI	Appropriateness of questioning techniques	3.65
28	SCR	Involves parents in learning of kids	3.62
29	HR	Participation in staff meetings	3.60
30	SCR	Holding parental conferences	3.50

Categories:
CI = Curriculum and Instruction
PD = Personal Development
SCR = School-Community Relations
HR = Human Relations

Question 4: Differences Between Teacher Interns' and Supervising Teachers' Perceptions Concerning Performance Effectiveness of Teacher Interns.
Ten tasks that received the highest mean ratings of performance effectiveness as rated by teacher interns were Human Relations (8), Curriculum and Instruction (1), and Personal Development (1). Similarly, the top 10 tasks, according to supervising teachers' ratings, were Human Relations (8), Curriculum and Instruction (1), and Personal Development (1). Of the top 10 tasks that received highest ratings on the two lists, seven tasks appeared on both the interns' and supervising teachers' lists. All the seven tasks were in the area of Human Relations. Although there were eight Human Relations, one Curriculum and Instruction, and one Personal Development task on both the interns' and supervising teachers' lists and that of the eight Human Relations tasks seven tasks were common to both lists, the tasks were ranked differently on the two lists.

The 10 tasks that received the lowest ratings of performance effectiveness by both groups show that nine of the tasks were the same on the two lists, although the rankings differed. However, the mean rating scores of supervising teachers were consistently higher than corresponding ratings of teacher interns on both the top 10 and bottom 10 tasks, indicating that supervising teachers felt that teacher interns were performing better than the interns themselves perceived.

Because there was a significant difference between the perceptions of teacher interns and their supervising teachers concerning the extent of the performance effectiveness of teacher interns, further analysis of the differences of the perceptions of the two groups in each task was conducted.

Those job dimensions for which differences between the perceptions of supervising teachers and teacher interns were statistically significant appear in Table 6.

According to Table 6, only five tasks produced statistically significant differences for paired means beyond the 0.05 level in a two-tailed test. Of these five tasks, two were among the bottom 10 and one was in the top 10 tasks according to both supervising teachers' and teacher interns' ratings. Three of the five task appearing in Table 6 were classified as "Human Relations" (HR) and one each as "School/Community Relations" (SCR) and "Personal Development" (PD). No "Curriculum and Instruction" (CI) task appeared among those that showed statistically significant differences in the extent of job performance as perceived by interns and their supervising teachers.

Generally, the supervising teachers' ratings of the extent of job performance of teacher interns were higher for all tasks rated than self ratings of teacher interns. This finding is at variance with most research studies including those by Ash (1980), Hubert and Dueck (1985), Johnston and Sackeney (1982), Lawler and Porter (1967), Levin (1980), Meyer (1980), and Thornton (1968, 1980) that concur that self ratings are generally inflated compared with counterposition appraisals. Only Heneman's (1974) study, which found that six mean self ratings out of nine were lower than supervisor ratings, provides support for the current finding. Several explanations of why self ratings were less lenient than and correlated to supervisor ratings are suggested, including the uniqueness of the research setting and degree of familiarity with the rating scales used in the study.

Most studies on self appraisals have been carried out in noneducational settings where subordinate-superordinate relationships are quite pronounced and hence tend to create a threatening atmosphere in the evaluation process. Second, the ultimate use to which data from self appraisals is put may influence the objectivity of these ratings. A ratee may be readily persuaded to inflate self ratings in order to provide a more positive image that promotes

Table 6
A t-test Analysis of the Perceptions of Supervising Teachers and Teacher Interns Concerning Extent of Job Performance Effectiveness by Teacher Interns (N=80)

*Category**	*Task*	*t Value*	*p Value***
HR	Consistence and fairness with students	2.54	0.013
SCR	Participation in community activities	3.96	0.002
HR	Participation in staff meetings	2.10	0.043
HR	Reflects and acts on supervisory advice	2.30	0.024
PD	Ingenuity and innovativeness	2.00	0.045

* Category:
HR = Human Relations
SCR = School/Community Relations
PD = Personal Development
** Probability value based on 2-tailed test of significance

favorable administrative decisions if such data were eventually used to influence administrative decisions than if they were used for purely research purposes as was the case in this study. Heneman (1974), whose findings are also supported by the present study, attributed lower self appraisal scores to "the fact that ratings were obtained under conditions in which the ratings could be used only for research, as opposed to evaluation and decision making purposes" (p. 642).

Finally, familiarity with rating scales and their expectations may lead to more realistic self-assessments. According to Thornton (1980), a clearer understanding of the meaning of the concepts to be measured increases the accuracy of interpreting the rating scales, which in turn leads to greater objectivity of assessments. Thus because the rating scale used in this investigation (Appendix B) was developed after those used by the Faculties of Education for teaching practice supervision as well as those used by cooperating schools for teacher interns' performance appraisals, teacher interns in this sample were likely to be familiar with the rating scale. Consequently, familiarity with and clearer conceptualization of the constructs and expectations of the rating scale by the research sample may have contributed to more realistic assessments as suggested by Thornton (1980).

Question 5: The Nature of Perceived Job Satisfaction of Teacher Interns and the Factors Contributing to Job Satisfaction.

The respondents' mean scores on the 21 Minnesota Satisfaction Questionnaire (MSQ) scales were compared with those of the MSQ normative data for teachers and it was found that the mean scores of teacher interns were consistently lower than those of the norm group on all MSQ facets including overall job satisfaction. In order to measure the degree of association between the perceptions of the sample and norm groups, a Spearman r was computed for paired rank ordered means of the two groups. A very high Spearman $r_s = 0.86$ was obtained, indicating a strong positive relationship between the perceptions of the two groups. This was evidence that there was relative consistency in the rankings of the two groups, further indicating that the two groups held similar preferences for MSQ job satisfaction facets. However, the two groups differed substantially in their rating of the extent of perceived job satisfaction as indicated in Table 7, which lists ordered deviations of sample values around the norm means for all MSQ facets.

According to Table 7, the range of deviations of sample values around the norm means was –10.47 (maximum) and –0.23 (minimum). Areas of greater discrepancy between the scores of the two groups were Security (–10.47), Compensation (–9.58) and Advancement (–8.36), followed by Authority, Social Service and Social Status (–2.58, –2.43, and –1.96 respectively). When sample means were rank ordered, Achievement, Creativity, Moral Values, Co-workers, and Supervision-Human Relations emerged as areas of greater job satisfaction, while Security, Compensation, Advancement, Authority and Social Status were areas of least job satisfaction. This finding also provides tentative support for Herzberg's Hygiene Theory because areas of greater and lesser satisfaction generally coincide with Herzberg's motivational and hygiene factors respectively (Herzberg, Mausuer, & Snyderman, 1959).

Table 7
A Comparison of Mean Scores for Job Satisfaction of Teacher Interns with Established MSQ Normative Data for Teachers (N=40)

MSQ Job Facet	*Sample Mean*	*Norm Mean*	*Deviation from Norm Mean*
Security	10.06	20.53	–10.47
Compensation	9.37	18.95	–9.58
Advancement	10.77	19.13	–8.36
Authority	16.37	18.95	–2.58
Social Service	19.65	20.08	–2.43
Social Status	16.82	18.78	–1.96
Creativity	19.90	21.83	–1.93
Working Condition	18.80	20.75	–1.87
Ability Utilization	19.25	21.08	–1.83
Recognition	17.65	19.46	–1.81
Variety	19.08	20.88	–1.80
Supervision-Human Relations	20.04	21.74	–1.70
Supervision-Technical	19.75	21.39	–1.64
Independence	18.88	20.52	–1.64
Co-workers	19.98	21.47	–1.49
Responsibility	19.48	20.90	–1.42
Activities	19.52	20.75	–1.23
Achievement	20.25	21.47	–1.22
Moral Values	20.23	21.42	–1.19
Company Policies and Practices	17.83	17.60	+0.23

Question 6. The Nature of the Relationship Between Role Clarity, Job Performance, and Job Satisfaction

Pearson product moment correlation analyses were computed to determine the nature of the relationships between self appraisals and supervisor appraisals with job satisfaction and that between role clarity with self and supervisor appraisals. All the correlation coefficients obtained for these relationships were positive, thus indicating the direction of the relationships. However, self appraisals produced larger and more statistically significant correlation coefficients ($p>0.05$) with both job satisfaction facets and role clarity dimensions than supervisor appraisals. This suggests that the relationships linking self appraisals to job satisfaction and role clarity were stronger than the corresponding relationships between supervisor appraisals and the latter variables. These results may indicate the degree to which the focal persons (teacher interns) valued their perceptions on job performance, which resulted in greater association between their perceptions with both job satisfaction and role clarity than the corresponding association between perceptions of supervising teachers with job satisfaction and role clarity. Thus the greater the importance or value attached to the perceptions, the stronger was the predictive relationship was between the predictor and criterion variables.

In order to determine the nature of the relationships linking self assessed performance, supervisor appraisals, role clarity, and job satisfaction and estab-

lish whether supervisor appraisals remained consistently weakly linked to the other research variables, a correlation matrix was computed for all four variables (Table 8). The results of the computation showed that supervisor appraisals failed to produce statistically significant coefficients of correlation with the other three research variables except for self appraisals. For these reasons, supervisor appraisals were left out of further analyses to determine the relationships linking role clarity, job performance, and job satisfaction.

At this point, the conceptual framework that was used to guide the study was utilized in interpreting the results of the correlation matrix concerning role clarity, job performance (self assessed), and job satisfaction. In order to explore further the relationships proposed in the conceptual framework, a partial correlation analysis was computed so that the indirect contributions of role clarity and self assessed performance to the proposed relationships linking the three research variables were established. The results of the partial correlation analysis showed that when role clarity was partialed out from both self assessed performance and job satisfaction, a partial correlation coefficient of 0.47 was obtained. This is only 0.085 smaller than the zero order correlation coefficient (0.555) between self assessed performance and job satisfaction. However, when self assessed performance was partialed out from the relationship between role clarity and job satisfaction, the correlation dramatically dropped from 0.388 to 0.19. This suggests that the relationship between role clarity and job satisfaction is mediated by self assessed performance.

The above findings provide support to the conceptual framework for this study, which proposed that role clarity (obverse of role ambiguity) is directly, but weakly, linked to job satisfaction as well as indirectly through job performance. However, the most significant finding of this study concerning the relationships linking the research variables was the emergence of self assessed performance as the mediating variable in the relationship between role clarity and job satisfaction. This finding further highlights the role played by self assessed performance in the role ambiguity-job performance-job satisfaction equation, a role that may have important implications to current teacher performance appraisal practices, which are dominated by supervisor appraisal systems.

Table 8
Pearson Product Moment Correlation Coefficients Matrix for Role Clarity, Self Appraisals, Supervisor Appraisals, and Job Satisfaction (N=80)

	Role Clarity	*Supervisor Appraisals*	*Self Appraisals*	*Job Satisfaction*
Role Clarity	1.000			
Supervisor Appraisals	0.193	1.00		
Self Appraisals	0.430*	0.403*	1.000	
Job Satisfaction	0.388*	0.230	0.555*	1.000

* value statistically significant at 0.05 level

Conclusions

Six major conclusions highlighting the findings of this study are enumerated below.

1. There was a high level of agreement between questionnaire and interview data indicating that overall the surveyed teacher interns generally experienced role clarity. Areas that contributed most to role clarity were concerned with "Role Responsibilities" while those concerned with "Performance Feedback" contributed least.
2. Teacher interns and their supervising teachers differed significantly in their perceptions of the extent of performance effectiveness of the interns over all 30 performance tasks. An examination of the mean ratings of the two groups over the 30 tasks showed that supervising teachers felt that teacher interns were performing more effectively than the interns perceived. The fact that the results of the study were to be used for research purposes only may have contributed to lower mean ratings of teacher interns.
3. The measures of association between the perceptions of supervising teachers and teacher interns concerning performance effectiveness were very high. The similarity in the ratings of the two groups may have been influenced by the fact that the groups were familiar with the rating scale used and that they had a clearer understanding of the meaning of the concepts to be measured, thus increasing the accuracy of interpreting the scale.
4. Teacher interns in this study experienced generally lower levels of job satisfaction, especially when compared with the Minnesota Satisfaction Questionnaire (MSQ) normative data for teachers. The nature of the program may have contributed to low feelings of job satisfaction among interns in that although they were fully qualified teachers in their own right, they did not enjoy the same conditions of service and other benefits as those enjoyed by other teachers.
5. There was a high measure of association between perceptions of the interns and norm group indicating that the two groups held similar preferences for MSQ job satisfaction facets.
6. The relationships between the research variables showed that self assessed performance was significantly linked to both role clarity and job satisfaction, while the relationships between supervisor assessed performance and the other variables were weak. In addition, self assessed performance mediated the relationship between role clarity and job satisfaction such that role clarity was linked directly, but weakly, to job satisfaction as well as indirectly via self assessed performance.

Implications

The findings of this study underline the importance of continued research concerning the concepts of self and supervisor assessed performance. Of particular importance is the role of self appraisals in relation to job satisfaction of teaching personnel. The close relationship between self appraisal and job satisfaction reported in this study raises the question of the utility of continued use of supervisor appraisals as the only teacher evaluation approach in education. The importance and value attached to self assessed performance as evidenced by the findings of this study suggests that educational administrators and educationists should explore further the possibility of making

greater use of self appraisal scales in education. Finally, areas of greater and lesser job satisfaction for teacher interns, as well as their perceptions concerning role clarity, provide an important source of information for developers of preservice programs and teacher internship in particular. The results suggest that job satisfaction may be greatly enhanced by restructuring and improving the nature of the job itself, while job dissatisfaction may be avoided by improving conditions of service of teachers. These two coincide with areas of perceived greater and lesser satisfaction as revealed by this study.

References

Alberta Education. (1985). *Initiation to teaching (Internship Project (ITP).* Edmonton : Author

Ash, R.A. (1980). Self-assessments of five types of typing ability. *Personnel Psychology, 33,* 273-282.

Beehr, T.A. (1976). Perceived situational moderators of the relationship between subjective role ambiguity and role strain. *Journal of Applied Psychology, 16,* 35-40.

Beehr, T.A., Walsh, J.T., & Taber, T.D. (1976). Relationship of stress to individually and organizationally valued states: Higher order needs as a moderator. *Journal of Applied Psychology, 61,* 41-47.

Burke, R.J., & Belcourt, M.L. (1974). Managerial role stress. *Journal of Business Administration, 5,* 55-68.

Cherniss, C. (1980). *Staff burnout: Job stress in human service organizations.* Beverly Hills, CA: Sage.

Cohen, A.R. (1959). Situational structure, self-esteem, and threat-oriented reaction to power. In D. Cartwright (Ed.), *Studies in social power.* Ann Arbor, MI : Research Center for Group Dynamics, Institute for Social Research.

Comadema, M.E. (1984). Brainstorming groups : Ambiguity tolerance, communication apprehension, task attraction and individual productivity. *Small Group Behavior, 15,* 251-264.

Farber, B.A. (1983). *Stress and burnout in the human service professions.* New York: Pergamon.

Friesen, D., Holdaway, E.A., & Rice, A.W. (1983). Satisfaction of school principals with the work. *Educational Administration Quarterly, 19,* 35-58.

French, W.L., Kass, F.E., & Rosenzweig, J.E. (1985). *Understanding human behavior in organizations.* New York: Harper & Row.

Greene, C.N. (1972). Relationships among role accuracy, compliance, performance evaluation, and satisfaction within managerial dyads. *Academy of Management Journal, 15,* 205-215.

Greene, C.N., & Organ, D.W. (1973). An evaluation of causal models linking the received role with job satisfaction. *Administrative Science Quarterly, 18,* 95-103.

Hamner, C., & Tosi, H. (1974). Relationship of role conflict and role ambiguity to job involvement measures. *Journal of Applied Psychology, 4,* 497-499.

Heneman, H.G. (1974). Comparisons of self and supervisor ratings of managerial performance. *Journal of Applied 59,* 638-642.

Herzberg, F., Mausuer, B., & Snyderman. (1959). *The motivation to work* (2nd ed.). New York: Wiley.

Holdaway, E.A. (1978). *Satisfaction of teachers in Alberta with their work and working conditions.* Edmonton, AB: Faculty of Education, University of Alberta.

Hewitson, W.T. (1975). *The professional satisfaction of beginning teachers.* Unpublished doctoral dissertation, University of Alberta.

House, R.J., & Rizzo, J.R. (1972). Role conflict and ambiguity as critical variables in a model of organizational behavior. *Organizational Behavior and Human Performance, 7,* 467-505.

Hubert, B.D., & Dueck, C.G. (1985). On-the-job training of assistant principals in selected tasks in the Calgary school system. *Alberta Journal of Educational Research, 31,* 270-287.

Johnston, J.M., & Sackeney, L.E. (1982). *Principals' classroom supervisory practices.* Saskatoon, SK: University of Saskatchewan, College of Education.

Kahn, R.L., Wolfe, D.M., Quinn, R.R., & Snoek, J.D. (1964). *Organizational stress: Studies in role conflict and ambiguity.* New York: Wiley.

Keller, R.T. (1975). Role conflict and ambiguity: Correlates with job satisfaction and values. *Personnel Psychology, 28,* 57-64.

Lawler, E.E., & Porter, L.W. (1967). The effect of performance on job satisfaction. *Industrial Relations, 7,* 20-28.

Levin, E.L. (1980). Introductory remarks for the symposium on organizational applications of self-appraisals and self assessment: Another look. *Personnel Psychology, 33,* 259-262.

Locke, E.A. (1970). Job satisfaction and job performance: A theoretical analysis. *Organizational Behavior and Human Performance, 5*, 484-500.
Locke, E.A. (1976). The nature and causes of job satisfaction. In M.D. Dunnette (Ed.), *Handbook of industrial and organizational psychology*. Chicago: Rand McNally.
Lortie, D.C. (1969). The balance of control and autonomy in elementary school teaching. In A. Etzioni (Ed.), *The semi-professions and their organization teachers, nurses, social workers*. New York: Free Press.
McLean, A. (1979). *Work stress*. Reading, MA: Addison-Wesley.
Meyer, H.H. (1980). Self-appraisal of the performance. *Personnel Psychology, 33*, 291-295.
Miles, R.H. (1976). A comparison of the relative impacts of role perceptions of ambiguity and conflict by role. *Academy of Management Journal, 19*, 25-35.
Miskel, C.G. (1982). Motivation in educational organizations. *Educational Administration Quarterly, 18*, 65-88.
Mouly, G.J. (1981). *The science of educational research*. New York: American Book Company.
Pigge, F.L., & Lovett, M.T. (1985). *Job performance and job satisfaction of beginning teachers*. (ERIC Document Reproduction Service No. ED 265 149)
Rizzo, J.R., House, R.J., & Lirtzman, S.I. (1970). Role conflict and role ambiguity in complex organizations. *Administrative Science Quarterly, 15*, 150-163.
Sergiovanni, T. (1967). Factors which affect satisfaction and dissatisfaction of teachers. *Journal of Educational Administration, 5*, 66-82.
Schmidt, G.L. (1976). Job satisfaction among secondary school administrators. *Educational Administration Quarterly, 12*, 68-86.
Schwab, D.P., & Cummings, L.L. (1970). Theories of performance and satisfaction: A review. *Industrial Relations, 9*, 409-430.
Schwab, R.L., & Iwanicki, E.F. (1982). Perceived role conflict, role ambiguity and teacher burnout. *Educational Administration Quarterly, 18*, 263-269.
Siegel, J.P., & Bowen, D. (1971). Satisfaction and performance: Causal relationships and moderating effects. *Journal of Vocational Behavior, 1*, 263-269.
Smith, E.E. (1957). The effects of clear and unclear role expectations on group productivity and defensiveness. *Journal of Abnormal and Social Psychology, 54*, 213-217.
Szilagyi, A. (1977). An empirical test of causal inference between role perceptions, satisfaction with work, performance and organizational level. *Personnel Psychology, 30*, 375-387.
Thornton, G.C. (1968). The relationship between supervisor and self-appraisal of executive performance. *Personnel Psychology 21*, 441-455.
Thornton, G.C. (1980). Psychometric perceptions of self-appraisal of job performance. *Personnel Psychology, 33*, 263-271.
Tosi, H. (1971). Organizational stress as moderator of the relationship between influence and role influence. *Academy of Management Journal, 14*, 7-20.
Van de Ven, A.H., & Ferry, D.L. (1980). *Measuring and asserting organizations*. New York: John Wiley and Sons.
Walsh, J.T., Taber T.D., & Beehr, T.A. (1980). An integrated model of perceived job characteristics. *Organizational Behavior and Human Performance, 25*, 252-267.
Weiss, D.J., Davis, R.V., England, G.W., & Lofquist, L.H. (1967). Manual for Minnesota Satisfaction Questionnaire. (Minnesota Studies in Vocational Rehabilitation: 22). University of Minnesota, Work Adjustment Project Industrial Relations Centre.
Wickstrom, R.A. (1973). Sources of teacher job satisfaction. *Canadian Administrator, 13*, 1-5.

Appendix A
Role Clarity Questionnaire Items

The role clarity questionnaire consisted of 16 items, which were rated using the following scale:

Always	Often	Frequently	Sometimes	Rarely	Never
6	5	4	3	2	1

I am certain about my responsibilities
I know exactly how much authority I have
School and jurisdiction policies and regulations are clear to me
I get clear and explicit directions and orders
I am certain as to who to report to in my job
I know exactly where to go for help
I get clear explanations of what to do
I know exactly how I am expected to relate to my colleagues
I know whether my supervisors will find my work acceptable
I have to ask others about what I am supposed to do
I have clear goals and objectives for my work
I am certain about how I will be evaluated
I receive feedback on how I do my work
I receive feedback when I do not expect it
I know exactly what my rights and privileges are
My terms and conditions of service have been made clear to me

Appendix B
Job Questionnaire Items

The job performance questionnaire consisted of 30 items which were rated by dyads of teacher interns and their supervising teachers using the following scale:

Poor	Fair	Good	Very Good	Exceptional
1	2	3	4	5

Preparation of long and short term plans
Designing appropriate objectives
Designing suitable student activities
Keeping accurate records
Cooperating with colleagues in lesson planning
Reinforcing students (e.g., appropriate feedback, praise, and so on)
Developing interesting and challenging learning activities

Classroom management and control
Responding to needs, aptitudes and learning styles of students
Using a variety of appropriate questioning techniques
Suitability of learning materials, illustrations and examples
Appropriateness of lesson introduction and closure
Ingenuity and innovativeness
Showing empathy for students
Student counseling
Showing consistence and fairness in dealing with students
Encouraging and valuing parental involvement in student work
Holding parental conferences
Student supervision
Participation in community activities
Participation in staff meetings
Maintaining good rapport with colleagues
Showing initiativeness
Providing good leadership (e.g., dependable, punctual, and so on)
Assisting in extracurricular activities
Reflecting and acting upon supervisory suggestions
Ability to make independent decisions
Maintaining good rapport with superiors
Developing own teaching approaches
Ability to make friendship with colleagues

For Contributors

Please prepare manuscripts in accordance with these guidelines.

Format

- Manuscript in triplicate.
- Typewritten or printed, double-spaced throughout.
- Abstract of approximately 100 words on a separate sheet.
- Manuscript not to exceed 6,500 words excluding graphics.
- Author's name and affiliation on a separate title page; title only on first page of manuscript to ensure anonymity in the review process.

Referencing

- Sources in parentheses after each reference (direct or otherwise), giving author's name (unless mentioned in text), year of publication, and page number(s).
- Direct quotes of 40 words or less in double quotation marks incorporated into text; those longer than 40 words in indented block format.
- All sources listed alphabetically at the end of the manuscript under the heading *References* using the APA style.
- Explanatory notes, numbered consecutively and marked in the text with superscript numerals, may appear before the References under the heading *Notes*; any citations in notes follow the same format as other references.

Graphics

- Each table and figure on a separate page at the end of the manuscript.
- Figures, charts, and diagrams in camera-ready form.
- Number tables in Arabic numerals and center over each heading, e.g., *Table 1*.
- Number figures in Arabic numerals before the caption at foot of each graph, e.g., *Figure 1: ...*
- Indicate placement of figures and tables in text, i.e., *Insert Table 3 here.*

Style

- *AJER*'s editorial style conforms closely to the *Publication Manual of the American Psychological Association.*
- Editorial changes may be made to manuscripts.
- For spelling, consult *Webster's New Collegiate Dictionary;* spelling in quoted material remains as in the original.

Diskettes

If manuscript is accepted for publication and has been prepared using a Macintosh, IBM, or IBM compatible microcomputer, please send diskette. Microsoft Word, MultiMate, WordPerfect, WordStar, Xerox Writer, XyWrite, and DisplayWrite 3 **only** are acceptable.
Important: Please indicate program used.

Submission to *AJER* implies the manuscript is not being considered for publication elsewhere.

AJER is indexed in the *Alberta Education Index, Canadian Education Index, Current Contents/Social and Behavioral Sciences, ERIC/Current Index to Journals in Education,* and *Social Science Citation Index*; appropriate articles are abstracted in *Cognitive Development Abstracts, Educational Administration Abstracts, Educational Technology Abstracts, Family Abstracts, Language Behavior Abstracts, Multicultural Education Abstracts, Psychological Abstracts, Research into Higher Education Abstracts, School Organization and Management Abstracts,* and *Sociology of Education Abstracts,* and reviewed by the *Journal for Research in Mathematics Education.*

If undelivered return to:

The Alberta Journal of Educational Research
4-116 Education North
University of Alberta
Edmonton, AB
T6G 2G5

Publications Mail Registration Number 1436